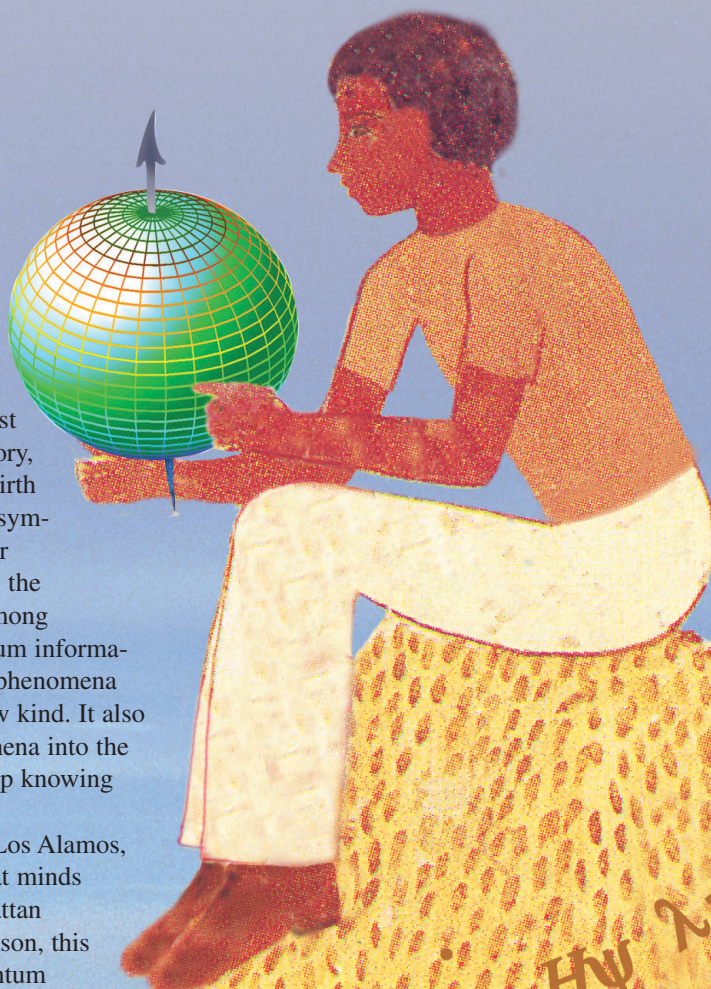


In the illustration at right, a young man holds a qubit in his hands. His perspective rests on knowledge accumulated over the last century in quantum physics, information theory, and computer science, the fields that gave birth to the concept of quantum information. He symbolizes the potential of this new resource for communication and computation, as well as the curiosity and excitement it has generated among young men and women. Research on quantum information holds the promise of making quantum phenomena subject to control and manipulation of a new kind. It also holds the promise of bringing these phenomena into the classroom, where young people will grow up knowing the quantum first hand.

Inspiration is derived in many ways. At Los Alamos, a sense of history and the legacy of the great minds who were leading participants in the Manhattan Project are a continuing source. For that reason, this volume about the Los Alamos effort in quantum information and quantum science opens with thought-provoking words from John Wheeler and Richard Feynman (see [pages vi–ix](#)). Both were Manhattan Project pioneers, and as discussed below, both have helped launch the field of quantum information science and renew interest in the foundations of quantum theory and measurement.



$$E = \hbar\omega \quad i\hbar\psi = H\psi \quad \lambda = h/p$$

$$|\langle ab \rangle - \langle ac \rangle| \leq 1 - \langle bc \rangle \quad S = -\text{Tr} \rho \ln \rho$$

$$\Delta x \Delta p \geq \hbar/2 \quad (\alpha|\uparrow\rangle + \beta|\downarrow\rangle)|0\rangle = \alpha|\uparrow\rangle|0\rangle + \beta|\downarrow\rangle|0\rangle$$

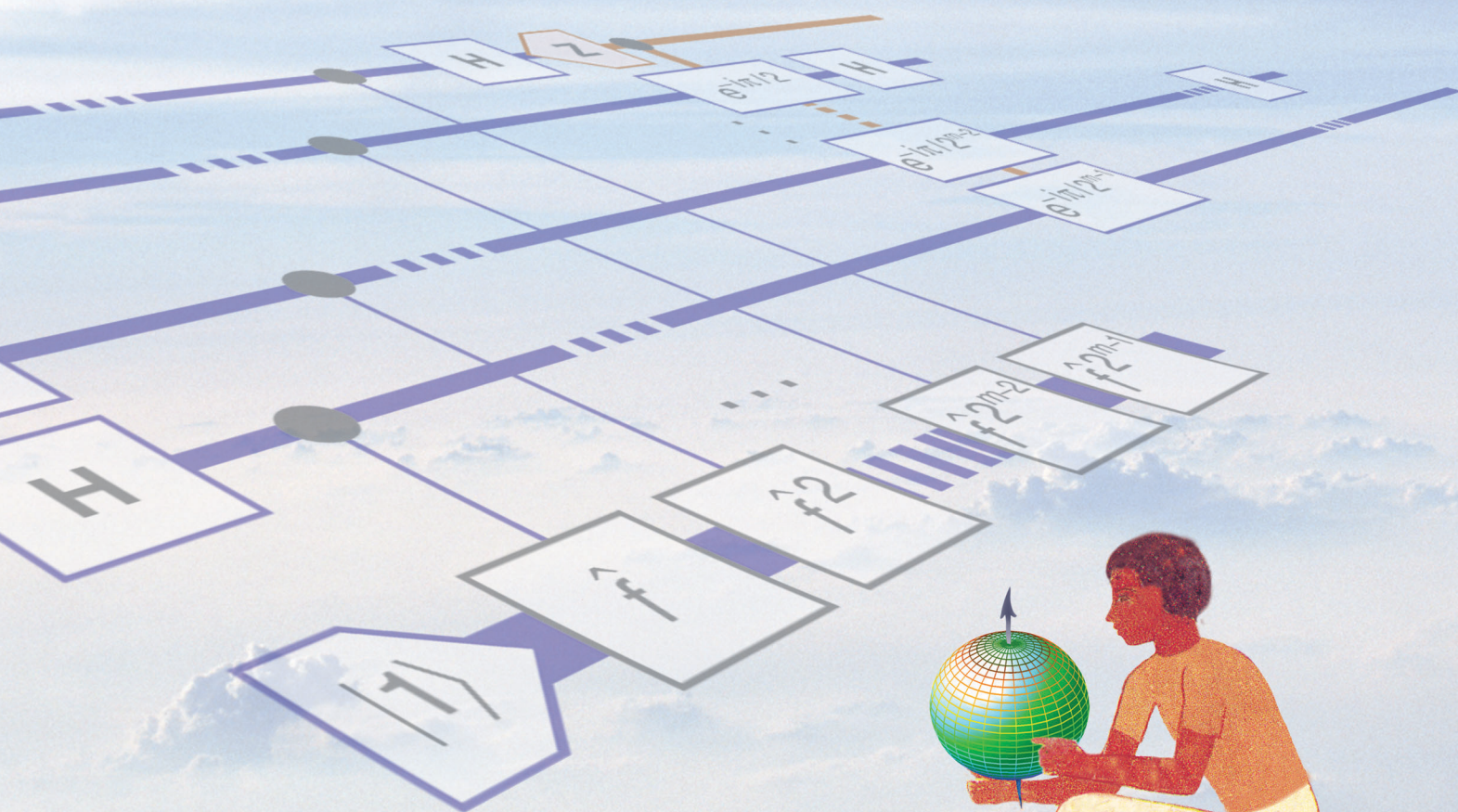
$$\text{cnot} = |0\rangle_A A|0\rangle + |1\rangle_A A|1\rangle (|0\rangle_B B|1\rangle + |1\rangle_B B|0\rangle)$$

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| \quad E = \hbar\omega$$

Quantum Information Processing

A hands-on primer

*Emanuel Knill, Raymond Laflamme, Howard N. Barnum, Diego A. Dalvit, Jacek J. Dziarmaga,
James E. Gubernatis, Leonid Gurvits, Gerardo Ortiz, Lorenza Viola, and Wojciech H. Zurek*



$$\alpha|0\rangle + \beta|1\rangle$$

Quantum information processing, Science of—The theoretical, experimental and technological areas covering the use of quantum mechanics for communication and computation
—Kluwer Encyclopedia of Mathematics, Supplement II

Research conducted in the last few decades has established that quantum information, or information based on quantum mechanics, has capabilities that exceed those of traditional “classical” information. For example, in communication, quantum information enables quantum cryptography, which is a method for communicating in secret. Secrecy is guaranteed because eavesdropping attempts necessarily disturb the exchanged quantum information without revealing the content of the communication. In computation, quantum information enables efficient simulation of quantum physics, a task for which general-purpose, efficient, classical algorithms are not known to exist. Quantum information also leads to efficient algorithms for factoring large numbers, which is believed to be difficult for classical computers. An efficient factoring algorithm would break the security of commonly used public-key cryptographic codes used for authenticating and securing Internet communications. Yet another application of quantum information improves the efficiency with which unstructured search problems can be solved. Quantum unstructured search may make it possible to solve significantly larger instances of optimization problems, such as the scheduling and traveling salesman problems.

Because of the capabilities of quantum information, the science of quantum information processing is now a prospering, interdisciplinary field focused on better understanding the possibilities and limitations of the underlying theory, on developing new applications of quantum information, and on physically realizing controllable quantum devices. The purpose of this primer is to provide an elementary introduction to quantum information processing (see Part II), and then to briefly explain how we hope to exploit the advantages of quantum information (see Part III). These two sections can be read independently. For reference, we have included a glossary of the main terms of quantum information (see [page 33](#)).

When we use the word “information,” we generally think of the things we can talk about, broadcast, write down, or otherwise record. Such records can exist in many forms, such as sound waves, electrical signals in a telephone wire, characters on paper, pit patterns on an optical disk, or magnetization on a computer hard disk. A crucial property of information is that it is fungible: It can be represented in many different physical forms and easily converted from one form to another without changing its meaning. In this sense, information is independent of the devices used to represent it but requires at least one physical representation in order to be useful.

We call the familiar information stored in today’s computers classical or deterministic to distinguish it from quantum information. It is no accident that classical information is the basis of all human knowledge. Any information passing through our senses is best modeled by classical discrete or continuous information. Therefore, when considering any other kind of information, we need to provide a method for extracting classically meaningful information. We begin by recalling the basic ideas of classical information in a way that illustrates the general procedure for building an information-processing theory.

Part I: Classical Information

The basic unit of classical deterministic information is the bit, an abstract entity or system that can be in one of the two states symbolized by 0 and 1. At this point, the symbols for the two states have no numeric meaning. That is why we have used a font different from that for the numbers 0 and 1. By making a clear distinction between the bit and its states, we emphasize that a bit should be physically realized as a system or device whose states correspond to the ideal bit's states. For example, if you are reading this primer on paper, the system used to realize a bit is a reserved location on the surface, and the state depends on the pattern of ink (0 or 1) in that location. In a computer, the device realizing a bit can be a combination of transistors and other integrated-circuit elements with the state of the bit determined by the distribution of charge.

In order to make use of information, it must be possible to manipulate (or process) the states of information units. The elementary operations that can be used for this purpose are called gates. Two one-bit gates are the **not** and **reset** gates. Applying the **not** gate to a bit has the effect of flipping the state of the bit. For example, if the initial state of the bit is 0, then the state after applying **not** is **not** (0) = 1. We can present the effect of the gate in the following form:

Initial State		Final State
0	→	not (0) = 1 , and
1	→	not (1) = 0 .

(1)

The **reset** gate sets the state to 0 regardless of the input:

Initial State		Final State
0	→	reset (0) = 0 , and
1	→	reset (1) = 0 .

(2)

By applying a combination of **not** and **reset** gates, one can transform the state of a bit in every possible way.

Information units can be combined to represent more information. Bits are typically combined into sequences. The states of such a sequence are symbolized by strings of state symbols for the constituent bits. For example, a two-bit sequence can be in one of the following four states: 00, 01, 10, and 11. The different bits are distinguished by their position in the sequence.

The one-bit gates can be applied to any bit in a sequence. For example, the **not** gate applied to the second bit of a three-bit sequence in the state 011 changes the state to 001.

One-bit gates act independently on each bit. To compute with multiple bits, we need gates whose action can correlate the states of two or more bits. One such gate is the **nand** (“not and”) gate, which acts on two bits in a bit sequence. Its effect is to set the state of the first bit to 0 if both the first and the second bit are 1; otherwise, it sets it to 1. Here is what happens when **nand** is applied to two consecutive bits:

Initial State		Final State
00	→	nand (00) = 10 ,
01	→	nand (01) = 11 ,
10	→	nand (10) = 10 , and
11	→	nand (11) = 01 .

(3)

The **nand** gate can be applied to any two bits in a sequence. For example, it can be applied to the fourth and second bits (in this order) of four bits, in which case the initial state 1101 is transformed to 1100, setting the fourth bit to 0.

Other operations on bit sequences include adding a new bit to the beginning (prepend) or end (append) of a sequence. The new bit is always initialized to 0. It is also possible to discard the first or last bit regardless of its state. Versions of these operations that are conditional on the state of another bit may also be used. An example is the conditional append operation: “If the k^{th} bit is in the state 0, then append a bit.”

The gates just introduced suffice for implementing arbitrary state transformations of a given bit sequence. Instructions for applying gates in a particular order are called a circuit. An important part of investigations in information processing is to determine the minimum resources required to perform information-processing tasks. For a given circuit, the two primary resources are the number of gates and the total number of bits used. The circuit complexity of a desired transformation is the minimum number of gates needed to implement it.

The model of computation defined by the ability to apply gates in a fixed sequence is called the circuit model. Classical computation extends the circuit model by providing a means for repeating blocks of instructions indefinitely or until a desired condition is achieved. In principle, it is possible to conceive of a general-purpose computer as a device that repeatedly applies the same circuit to the beginnings of several bit sequences. In this article, we take for granted a traditional programmable computer based on classical information. Thus, a quantum algorithm is a program written for such a computer with additional instructions for applying gates to quantum information. The computational power of this model is equivalent to that of other general-purpose models of quantum computation, such as quantum Turing machines (Yao 1993).

For an introduction to algorithms and their analysis, refer to Thomas Cormen et al. (1990). Christos Papadimitriou wrote (1994) a useful textbook on computational complexity with an introduction to classical computation and computational machine models.

Part II: Quantum Information

The foundations of an information-processing theory can be constructed by the procedure we followed in the previous section:

1. Define the basic unit of information.
2. Give the means for processing one unit.
3. Describe how multiple units can be combined.
4. Give the means for processing multiple units.
5. Show how to convert the content of any of the extant units to classical information.

Note that the last step was not required for classical information processing.

In this section, we follow the general procedure for defining an information-processing theory to introduce quantum information processing. A simple example that exhibits the advantages of quantum information is given in the section “The Parity Problem” on [page 21](#). A version of the quantum factoring algorithm is described immediately following this article in “From Factoring to Phase Estimation” on [page 38](#).

The Quantum Bit

The fundamental resource and basic unit of quantum information is the quantum bit (qubit), which behaves like a classical bit enhanced by the superposition principle (see discussion in this section). From a physical point of view, a qubit is represented by an ideal two-state quantum system. Examples of such systems include photons (vertical and horizontal polarization), electrons and other spin-1/2 systems (spin-up and -down), and systems defined by two energy levels of atoms or ions. From the beginning, the two-state system played a central role in studies of quantum mechanics. It is the simplest quantum system, and in principle, all other quantum systems can be modeled in the state space of collections of qubits.

From the information-processing point of view, a qubit’s state space contains the two “logical,” or computational, states $|0\rangle$ and $|1\rangle$. The so-called “ket” notation for these states was introduced by Paul Dirac, and its variations are widely used in quantum physics. One can think of the pair of symbols $|$ and \rangle as representing the qubit system. Their content specifies a state for the system. In this context, 0 and 1 are system-independent state labels. When, say, 0 is placed within the ket, the resulting expression $|0\rangle$ represents the corresponding state of a specific qubit.

The initial state of a qubit is always one of the logical states. Using operations to be introduced later, we can obtain states that are superpositions of the logical states. Superpositions can be expressed as sums $\alpha|0\rangle + \beta|1\rangle$ over the logical states with complex coefficients. The complex numbers α and β are the amplitudes of the superposition. The existence of such superpositions of distinguishable states of quantum systems is one of the basic tenets of quantum theory and is called the superposition principle. Another way of writing a general superposition is as a vector:

$$\alpha|0\rangle + \beta|1\rangle \leftrightarrow \begin{pmatrix} \alpha \\ \beta \end{pmatrix}, \quad (4)$$

where the two-sided arrow is used to denote the correspondence between expressions that mean the same thing.

The qubit states that are superpositions of the logical states are called pure states: A superposition $\alpha|0\rangle + \beta|1\rangle$ is a pure state if the corresponding vector has length 1, that

is, $|\alpha|^2 + |\beta|^2 = 1$. Such a superposition or vector is said to be normalized. (For a complex number given by $\gamma = x + iy$, one can evaluate $|\gamma|^2 = x^2 + y^2$. Here, x and y are the real and imaginary part of γ , and the symbol i is a square root of -1 , that is, $i^2 = -1$. The conjugate of γ is $\bar{\gamma} = x - iy$. Thus, $|\bar{\gamma}|^2 = \gamma\bar{\gamma}$. Here are a few examples of states given in both the ket and vector notation:

$$|\psi_1\rangle = \frac{(|0\rangle + |1\rangle)}{\sqrt{2}} \leftrightarrow \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \quad (5)$$

$$|\psi_2\rangle = \frac{3}{5}|0\rangle - \frac{4}{5}|1\rangle \leftrightarrow \begin{pmatrix} 3/5 \\ -4/5 \end{pmatrix}, \quad \text{and} \quad (6)$$

$$|\psi_3\rangle = \frac{i3}{5}|0\rangle - \frac{i4}{5}|1\rangle \leftrightarrow \begin{pmatrix} i3/5 \\ -i4/5 \end{pmatrix}. \quad (7)$$

The state $|\psi_3\rangle$ is obtained from $|\psi_2\rangle$ by multiplication with i . It turns out that two states cannot be distinguished if one of them is obtained by multiplying the other by a phase $e^{i\theta}$. Note how we have generalized the ket notation by introducing expressions such as $|\psi\rangle$ for arbitrary states.

The superposition principle for quantum information means that we can have states that are sums of logical states with complex coefficients. There is another, more familiar type of information, whose states are combinations of logical states. The basic unit of this type of information is the probabilistic bit (pbit). Intuitively, a pbit can be thought of as representing the as-yet-undetermined outcome of a coin flip. Since we need the idea of probability to understand how quantum information converts to classical information, we briefly introduce pbits.

A pbit's state space is a probability distribution over the states of a bit. One very explicit way to symbolize such a state is by using the expression $\{p:0, (1-p):1\}$, which means that the pbit has probability p of being 0 and $1-p$ of being 1. Thus, a state of a pbit is a probabilistic combination of the two logical states, where the coefficients are nonnegative real numbers summing to 1. A typical example is the unbiased coin in the process of being flipped. If tail and head represent 0 and 1, respectively, the coin's state is $\{1/2:0, 1/2:1\}$. After the outcome of the flip is known, the state collapses to one of the logical states 0 and 1. In this way, a pbit is converted to a classical bit. If the pbit is probabilistically correlated with other pbits, the collapse associated with learning the pbit's logical state changes the overall probability distribution by a process called conditioning on the outcome.

A consequence of the conditioning process is that we never actually "see" a probability distribution. We only see classical deterministic bit states. According to the frequency interpretation of probabilities, the original probability distribution can only be inferred after one looks at many independent pbits in the same state $\{p:0, (1-p):1\}$: In the limit of infinitely many pbits, p is given by the fraction of pbits seen to be in the state 0. As we will explain, we can never see a general qubit state either. For qubits, there is a process analogous to conditioning. It is called measurement and converts qubit states to classical information.

Information processing with pbits has many advantages over deterministic information processing with bits. One advantage is that algorithms are often much easier to design and

analyze if they are probabilistic. Examples include many optimization and physics simulation algorithms. In some cases, the best available probabilistic algorithm is more efficient than any known deterministic algorithm. An example is an algorithm for determining whether a number is prime or not. It is not known whether every probabilistic algorithm can be derandomized efficiently. There are important communication problems that can be solved probabilistically but not deterministically. For a survey of these algorithms, see Rajiv Gupta (1994a).

What is the difference between bits, pbits, and qubits? One way to visualize the difference and see the enrichment provided by pbits and qubits is shown in Figure 1.

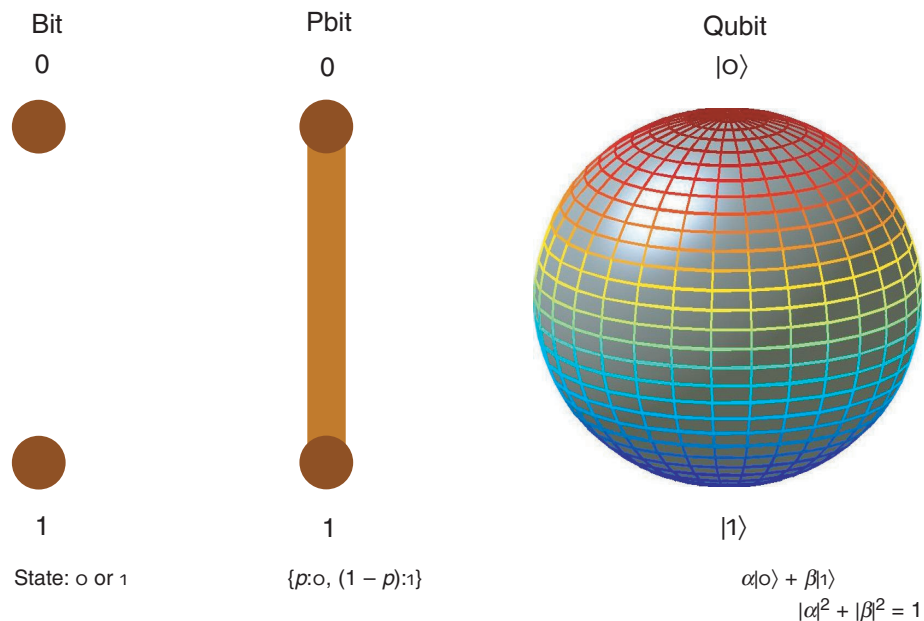


Figure 1. Comparing State Spaces of Different Information Units

The states of a bit correspond to two points. The states of a pbit can be thought of as convex combinations of a bit's states and therefore can be visualized as lying on the line connecting the two bit states. A qubit's pure states correspond to the surface of the unit sphere in three dimensions, where the logical states correspond to the poles. This representation of qubit states is called the Bloch sphere. The explicit correspondence is discussed at the end of the section "Mixtures and Density Operators." Also refer to the definition and use of the Bloch sphere in the article "NMR and Quantum Information Processing" on [page 226](#). There, the correspondence between the pure states and the sphere is physically motivated and comes from a way of viewing a spin-1/2 system as a small quantum magnet. Intuitively, a state is determined by the direction of the north pole of the magnet.

Processing One Qubit

The quantum version of the **not** gate for bits exchanges the two logical states; that is, using ket notation,

$$\text{not}(\alpha|0\rangle + \beta|1\rangle) = \alpha|1\rangle + \beta|0\rangle = \beta|0\rangle + \alpha|1\rangle. \quad (8)$$

In vector notation, this equation becomes

$$\mathbf{not} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \beta \\ \alpha \end{pmatrix}. \quad (9)$$

Another way of expressing the effect of **not** is by multiplying the vector by a matrix representing **not**,

$$\mathbf{not} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \beta \\ \alpha \end{pmatrix}, \quad (10)$$

so that we can identify the action of **not** with the matrix

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

An even simpler gate is the one that does nothing. We call it the **noop** gate, and its matrix form is the identity matrix, as shown in the following equation:

$$\mathbf{noop} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}. \quad (11)$$

The **noop** and **not** gates are reversible. In other words, we can undo their actions by applying other gates. For example, the action of the **not** gate can be undone by another **not** gate. The action of every reversible quantum gate can be represented by matrix multiplication, where the matrix has the additional property of preserving vector lengths. Such matrices are called unitary and are characterized by the equation $A^\dagger A = \mathbb{1}$, where A^\dagger is the conjugate transpose of A and $\mathbb{1}$ is the identity matrix. (The conjugate transpose of a matrix is computed by flipping that matrix across the main diagonal and conjugating the complex numbers). For gates represented by a matrices, the unitarity condition is necessary and sufficient for ensuring that pure states get mapped to pure states.

Because qubit states can be represented as points on a sphere, reversible one-qubit gates can be thought of as rotations of the Bloch sphere. This is why such quantum gates are often called rotations. As explained in detail on [page 232](#) in the article “NMR and Quantum Information Processing”, rotations around the x -, y -, and z -axis are in a sense generated by the three Pauli matrices

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \text{and} \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (12)$$

each of which represents a one-qubit gate. For example, a rotation around the x -axis by an angle ϕ is given by $e^{-i\sigma_x\phi/2} = \cos(\phi/2)\mathbb{1} - i\sin(\phi/2)\sigma_x$. To obtain this identity, one can use the power series for e^A , $e^A = \sum_{k=0}^{\infty} (1/k!)A^k$, and exploit the fact that $\sigma_x^2 = \mathbb{1}$ to simplify the expression. Here are some gates that can be defined with the help of rotations:

$$90^\circ \text{ } x\text{-rotation: } \mathbf{rotx}_{90^\circ} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ -i & 1 \end{pmatrix},$$

$$90^\circ \text{ } y\text{-rotation: } \mathbf{roty}_{90^\circ} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix},$$

$$\phi \text{ } z\text{-rotation: } \mathbf{rotx}_\phi = \begin{pmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{pmatrix}, \text{ and}$$

$$\text{Hadamard gate: } \mathbf{H} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \quad (13)$$

The rotation gates often show up in controlling spins or ions with radio-frequency pulses or lasers. The Hadamard gate is used primarily by quantum programmers. It can be expressed as a product of a 90° y -rotation and σ_z .

To check directly that the rotation gates are reversible, one can determine their inverses. In this case and as expected, the inverse of a rotation is the rotation around the same axis in the opposite direction. For example, the inverses of the \mathbf{roty}_{90° and \mathbf{rotx}_ϕ gates are given by

$$\mathbf{roty}_{-90^\circ} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \text{ and } \mathbf{rotx}_{-\phi} = \begin{pmatrix} e^{i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{pmatrix}. \quad (14)$$

Another useful property of the rotation gates is that the angles add when rotations are applied around the same axis. For example, $\mathbf{rotx}_\phi \mathbf{rotx}_\theta = \mathbf{rotx}_{\phi+\theta}$.

The Bra-Ket Notation for Logic Gates. The ket notation can be extended so that we can write gates in a compact form that readily generalizes to multiple qubits. To do so, we have to introduce expressions such as $\langle\psi| = \alpha\langle 0| + \beta\langle 1|$. This is called the “bra” notation. The terminology comes from the term “bracket.” The bra is the left, and the ket is the right part of a matched pair of brackets. From the vector point of view, $\langle\psi|$ corresponds to the row vector (α, β) . Note that a column vector multiplied by a row vector yields a matrix. In the bra-ket notation, this corresponds to multiplying a ket $|\psi\rangle$ by a bra $\langle\phi|$, written as $|\psi\rangle\langle\phi|$. Since this represents an operator on states, we expect to be able to compute the effect of $|\psi\rangle\langle\phi|$ on a state $|\phi\rangle$ by forming the product. To be able to evaluate such products with one-qubit bras and kets, we need the following two rules: distributivity and inner-product evaluation.

Distributivity

You can rewrite sums and products using distributivity. For example,

$$\left(\frac{3}{5}\langle 0| + \frac{4}{5}\langle 1| \right) i|1\rangle = \frac{i3}{5}\langle 0||1\rangle + \frac{i4}{5}\langle 1||1\rangle. \quad (15)$$

Observe that we can combine the amplitudes of terms, but we cannot rearrange the order of the bras and kets in a product.

Inner-Product Evaluation

The product of a logical bra and a logical ket is evaluated according to the identities

$$\begin{aligned}\langle 0|0\rangle &= 1 \\ \langle 0|1\rangle &= 0, \\ \langle 1|0\rangle &= 0, \text{ and} \\ \langle 1|1\rangle &= 1.\end{aligned}\tag{16}$$

It follows that for logical states, if a bra multiplies a ket, the result cancels unless the states match, in which case the answer is 1. Applying inner-product evaluation to Equation (15) results in

$$\frac{i3}{5}\langle 0|1\rangle + \frac{i4}{5}\langle 1|1\rangle = \frac{i3}{5}0 + \frac{i4}{5}1 = \frac{i4}{5}.\tag{17}$$

To simplify the notation, we can omit one of the two vertical bars in products such as $\langle a|b\rangle$ and write $\langle a|b\rangle$.

To understand inner-product evaluation, think of the expressions as products of row and column vectors. For example,

$$\langle 0|1\rangle \leftrightarrow \begin{pmatrix} 1 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0.\tag{18}$$

That is, as vectors, the two states $|0\rangle$ and $|1\rangle$ are orthogonal. In general, if $|\phi\rangle$ and $|\psi\rangle$ are states, then $\langle\phi|\psi\rangle$ is the inner product, or “overlap,” of the two states. In the expression for the overlap, we compute $\langle\phi|$ from $|\phi\rangle = \bar{\alpha}|0\rangle + \bar{\beta}|1\rangle$ by conjugating the coefficients and converting the logical kets to bras: $\langle\phi| = \alpha\langle 0| + \beta\langle 1|$. In the vector representation, this is the conjugate transpose of the column vector for $|\phi\rangle$, so the inner product agrees with the usual one. Two states are orthogonal if their overlap is zero. We write $|\phi\rangle^\dagger = \langle\phi|$ and $\langle\phi|^\dagger = |\phi\rangle$.

Every linear operator on states can be expressed with the bra-ket notation. For example, the bra-ket expression for the **noop** gate is **noop** = $|0\rangle\langle 0| + |1\rangle\langle 1|$. To apply **noop** to a qubit, you multiply its state on the left by the bra-ket expression

$$\begin{aligned}\text{noop}(\alpha|0\rangle + \beta|1\rangle) &= (|0\rangle\langle 0| + |1\rangle\langle 1|)(\alpha|0\rangle + \beta|1\rangle) \\ &= |0\rangle\langle 0|(\alpha|0\rangle + \beta|1\rangle) + |1\rangle\langle 1|(\alpha|0\rangle + \beta|1\rangle) \\ &= \alpha|0\rangle\langle 0|0\rangle + \beta|0\rangle\langle 0|1\rangle + \alpha|1\rangle\langle 1|0\rangle + \beta|1\rangle\langle 1|1\rangle \\ &= \alpha|0\rangle 1 + \beta|0\rangle 0 + \alpha|1\rangle 0 + \beta|1\rangle 1 \\ &= \alpha|0\rangle + \beta|1\rangle.\end{aligned}\tag{19}$$

One way to think about an operator such as $|a\rangle\langle b|$ is to notice that, when it is used to operate on a ket expression, the $\langle b|$ picks out the matching kets in the state, which are

then changed to $|a\rangle$. For example, we can write the **not** operation as **not** = $|0\rangle\langle 1| + |1\rangle\langle 0|$.

The coefficients of the $|a\rangle\langle b|$ in a bra-ket representation of a gate correspond to matrix entries in the matrix representation. The relationship is defined by

$$\alpha_{00}|0\rangle\langle 0| + \alpha_{01}|0\rangle\langle 1| + \alpha_{10}|1\rangle\langle 0| + \alpha_{11}|1\rangle\langle 1| \leftrightarrow \begin{pmatrix} \alpha_{00} & \alpha_{01} \\ \alpha_{10} & \alpha_{11} \end{pmatrix}. \quad (20)$$

Two Quantum Bits

Some states of two quantum bits can be symbolized by the juxtaposition (or multiplication) of the states of each quantum bit. In particular, the four logical states $|0\rangle|0\rangle$, $|0\rangle|1\rangle$, $|1\rangle|0\rangle$, and $|1\rangle|1\rangle$ are acceptable pure states for two quantum bits. In these expressions, we have distinguished the qubits by position (first or second). It is easier to manipulate state expressions if we explicitly name the qubits, say, A and B. We can then distinguish the kets by writing, for example, $|\psi\rangle_A$ for a state of qubit A. Now, the state $|0\rangle|1\rangle$ can be written with explicit qubit names (or labels) as

$$|0\rangle_A |1\rangle_B = |1\rangle_B |0\rangle_A = |01\rangle_{AB} = |10\rangle_{BA}. \quad (21)$$

Having explicit labels allows us to unambiguously reorder the states in a product of states belonging to different qubits. We say that kets for different qubits “commute.”

So far, we have seen four states of two qubits, which are the logical states that correspond to the states of two bits. As in the case of one qubit, we can use the superposition principle to get all the other pure states. Each state of two qubits is therefore of the form

$$\alpha|00\rangle_{AB} + \beta|01\rangle_{AB} + \gamma|10\rangle_{AB} + \delta|11\rangle_{AB}, \quad (22)$$

where α , β , γ , and δ are complex numbers. Again, there is a column vector form for the state,

$$\begin{pmatrix} \alpha \\ \beta \\ \gamma \\ \delta \end{pmatrix}, \quad (23)$$

and this vector has to be of unit length, that is, $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$. When using the vector form for qubit states, one has to be careful about the convention used for ordering the coefficients.

Other examples of two-qubit states in ket notation are the following:

$$\begin{aligned}
|\psi_1\rangle_{AB} &= \frac{1}{\sqrt{2}}(|0\rangle_A + |1\rangle_A)|1\rangle_B \ . \\
|\psi_2\rangle_{AB} &= \frac{1}{\sqrt{2}}(|0\rangle_A - |1\rangle_A) \frac{1}{\sqrt{2}}(|0\rangle_B + i|1\rangle_B) \\
&= \frac{1}{2}(|00\rangle_{AB} + i|01\rangle_{AB} - |10\rangle_{AB} - i|11\rangle_{AB}) \ . \\
|\psi_3\rangle_{AB} &= \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB}) \ . \\
|\psi_4\rangle_{AB} &= \frac{1}{\sqrt{2}}(|01\rangle_{AB} - |10\rangle_{AB}) \ .
\end{aligned} \tag{24}$$

The first two of these states have the special property that they can be written as a product $|\phi_1\rangle_A|\phi_2\rangle_B$ of a state of qubit A and a state of qubit B. The second expression for $|\psi_2\rangle$ shows that the product decomposition is not always easy to see. Such states are called product states. The last two states, $|\psi_3\rangle_{AB}$ and $|\psi_4\rangle_{AB}$, are two of the famous Bell states. They have no such representation as a product of independent states of each qubit. They are said to be entangled because they contain a uniquely quantum correlation between the two qubit systems. Pbits can also have correlations that cannot be decomposed into product states, but the entangled states have additional properties that make them very useful. For example, if Alice and Bob each have one of the qubits that together are in the state $|\psi_3\rangle_{AB}$, they can use them to create a secret bit for encrypting their digital communications (see the article “Quantum State Entanglement” on [page 52](#)).

Processing Two Qubits

The simplest way of modifying the state of two qubits is to apply one of the one-qubit gates. If the gates are expressed in the bra-ket notation, all we need to do is add qubit labels so that we know which qubit each bra or ket belongs to. For example, the **not** gate for qubit B is written as

$$\mathbf{not}^{(B)} = |0\rangle_B^B \langle 1| + |1\rangle_B^B \langle 0| \ . \tag{25}$$

The labels for bra expressions occur as left superscripts. To apply expressions like this to states, we need one more rule, namely, commutation.

Commutation

Kets and bras with different labels can be interchanged in products (they commute). This property is demonstrated by the following example:

$$\begin{aligned}
(|0\rangle_B^B \langle 1|) |01\rangle_{AB} &= |0\rangle_B^B \langle 1| |0\rangle_A |1\rangle_B \\
&= |0\rangle_A |0\rangle_B^B \langle 1| |1\rangle_B \\
&= |0\rangle_A |0\rangle_B^B \langle 1| |1\rangle_B \\
&= |0\rangle_A |0\rangle_B = |00\rangle_{AB} \ .
\end{aligned} \tag{26}$$

Note that we cannot merge the two vertical bars in expressions such as ${}^B\langle 1||0\rangle_A$ because the two terms belong to different qubits. The bars can only be merged when the expression is an inner product, which requires that the two terms belong to the same qubit.

With the rules for bra-ket expressions in hand, we can apply the **not** gate to one of our Bell states to see how it acts:

$$\begin{aligned}
\mathbf{not}^{(B)} \frac{1}{\sqrt{2}} (|00\rangle_{AB} + |11\rangle_{AB}) &= \left(|0\rangle_B^B \langle 1| + |1\rangle_B^B \langle 0| \right) \frac{1}{\sqrt{2}} (|00\rangle_{AB} + |11\rangle_{AB}) \\
&= \frac{1}{\sqrt{2}} \left(|0\rangle_B^B \langle 1| (|00\rangle_{AB} + |11\rangle_{AB}) + |1\rangle_B^B \langle 0| (|00\rangle_{AB} + |11\rangle_{AB}) \right) \\
&= \frac{1}{\sqrt{2}} \left(|0\rangle_B^B \langle 1||00\rangle_{AB} + |0\rangle_B^B \langle 1||11\rangle_{AB} + |1\rangle_B^B \langle 0||00\rangle_{AB} + |1\rangle_B^B \langle 0||11\rangle_{AB} \right) \\
&= \frac{1}{\sqrt{2}} \left(|0\rangle_A |0\rangle_B^B \langle 1||0\rangle_B + |1\rangle_A |0\rangle_B^B \langle 1||1\rangle_B + |0\rangle_A |1\rangle_B^B \langle 0||0\rangle_B + |1\rangle_A |1\rangle_B^B \langle 0||1\rangle_B \right) \\
&= \frac{1}{\sqrt{2}} \left(|0\rangle_A |0\rangle_B 0 + |1\rangle_A |0\rangle_B 1 + |0\rangle_A |1\rangle_B 1 + |1\rangle_A |1\rangle_B 0 \right) \\
&= \frac{1}{\sqrt{2}} \left(|1\rangle_A |0\rangle_B + |0\rangle_A |1\rangle_B \right) = \frac{1}{\sqrt{2}} (|01\rangle_{AB} + |10\rangle_{AB}) .
\end{aligned} \tag{27}$$

The effect of the gate was to flip the state symbols for qubit B, which results in another Bell state.

The gate $\mathbf{not}^{(B)}$ can also be written as a 4×4 matrix acting on the vector representation of a two-qubit state. However, the relationship between this matrix and the one-qubit matrix is not as obvious as for the bra-ket expression. The matrix is

$$\mathbf{not}^{(B)} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \tag{28}$$

which swaps the top two and bottom two entries of a state vector.

One way to see the relationship between the one- and two-qubit representations of the gate $\mathbf{not}^{(B)}$ is to notice that because the **noop** gate acts as the identity and because we can act on different qubits independently, $\mathbf{noop}^{(A)} \mathbf{not}^{(B)} \equiv \mathbf{not}^{(B)}$. The matrix for $\mathbf{not}^{(B)}$ can be expressed as a Kronecker product (\otimes) of the matrices for **noop** and **not**:

$$\begin{aligned}
\text{noop}^{(A)} \text{not}^{(B)} &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \otimes \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & 0 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ 0 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} & 1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{pmatrix} \\
&= \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.
\end{aligned} \tag{29}$$

The Kronecker product of two matrices expands the first matrix by multiplying each entry by the second matrix. A disadvantage of the matrix representation of quantum gates is that it depends on the number and order of the qubits. However, it is often easier to visualize what the operation does by writing down the corresponding matrix.

One cannot do much with one-bit classical gates. Similarly, the utility of one-qubit gates is limited. In particular, it is not possible to obtain a Bell state starting from $|00\rangle_{AB}$ or any other product state. We therefore need to introduce at least one two-qubit gate not expressible as the product of two one-qubit gates. The best-known such gate is the controlled-not (**cnot**) gate. Its action can be described by the statement, “if the first bit is 1, flip the second bit; otherwise, do nothing.” The bra-ket and matrix representations for this action are

$$\begin{aligned}
\text{cnot}^{(AB)} &= |0\rangle_A^A \langle 0| + |1\rangle_A^A \langle 1| \left(|0\rangle_B^B \langle 1| + |1\rangle_B^B \langle 0| \right) \\
&= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.
\end{aligned} \tag{30}$$

The **cnot** gate is reversible because its action is undone if a second **cnot** is applied. This outcome is easy to see by computing the square of the matrix for **cnot**, which yields the identity matrix. As an exercise in manipulating bras and kets, let us calculate the product of two **cnot** gates by using the bra-ket representation:

$$\begin{aligned}
\text{cnot}^{(AB)} \text{cnot}^{(AB)} &= \left(|0\rangle_A^A \langle 0| + |1\rangle_A^A \langle 1| \left(|0\rangle_B^B \langle 1| + |1\rangle_B^B \langle 0| \right) \right) \\
&\quad \times \left(|0\rangle_A^A \langle 0| + |1\rangle_A^A \langle 1| \left(|0\rangle_B^B \langle 1| + |1\rangle_B^B \langle 0| \right) \right). \tag{31}
\end{aligned}$$

The first step is to expand this expression by multiplying out. Expressions such as $|\mathbf{o}\rangle_A^A \langle \mathbf{o} | | \mathbf{1} \rangle_A^A \langle \mathbf{1} |$ cancel because of the inner-product evaluation rule ${}^A \langle \mathbf{o} | \mathbf{1} \rangle_A = 0$. One can also reorder bras and kets with different labels and rewrite $|\mathbf{o}\rangle_A^A \langle \mathbf{o} | | \mathbf{o} \rangle_A^A \langle \mathbf{o} | = |\mathbf{o}\rangle_A^A \langle \mathbf{o} |$ to get

$$\begin{aligned}
 \mathbf{cnot}^{(AB)} \mathbf{cnot}^{(AB)} &= |\mathbf{o}\rangle_A^A \langle \mathbf{o} | + |\mathbf{1}\rangle_A^A \langle \mathbf{1} | \left(|\mathbf{o}\rangle_B^B \langle \mathbf{1} | + |\mathbf{1}\rangle_B^B \langle \mathbf{o} | \right) \left(|\mathbf{o}\rangle_B^B \langle \mathbf{1} | + |\mathbf{1}\rangle_B^B \langle \mathbf{o} | \right) \\
 &= |\mathbf{o}\rangle_A^A \langle \mathbf{o} | + |\mathbf{1}\rangle_A^A \langle \mathbf{1} | \left(|\mathbf{o}\rangle_B^B \langle \mathbf{o} | + |\mathbf{1}\rangle_B^B \langle \mathbf{1} | \right) \\
 &= |\mathbf{o}\rangle_A^A \langle \mathbf{o} | + |\mathbf{1}\rangle_A^A \langle \mathbf{1} | \mathbf{noop}^{(B)} \\
 &\equiv |\mathbf{o}\rangle_A^A \langle \mathbf{o} | + |\mathbf{1}\rangle_A^A \langle \mathbf{1} | \\
 &= \mathbf{noop}^{(A)} \\
 &\equiv 1 .
 \end{aligned} \tag{32}$$

Here we used the fact that, when the bra-ket expression for **noop** is applied to the ket expression for a state, it acts the same as (here denoted by the symbol \equiv) multiplication by the number 1.

Using Many Quantum Bits

To use more than two, say, five qubits, we can just start with the state $|\mathbf{o}\rangle_A |\mathbf{o}\rangle_B |\mathbf{o}\rangle_C |\mathbf{o}\rangle_D |\mathbf{o}\rangle_E$ and apply gates to any one or two of these qubits. For example, $\mathbf{cnot}^{(DB)}$ applies the **cnot** operation from qubit D to qubit B. Note that the order of D and B in the label for the **cnot** operation matters. In the bra-ket notation, we simply multiply the state with the bra-ket form of $\mathbf{cnot}^{(DB)}$ from the left. One can express everything in terms of matrices and vectors, but now the vectors have length $2^5 = 32$, and the Kronecker product expression for $\mathbf{cnot}^{(DB)}$ requires some reordering to enable inserting the operation so as to act on the intended qubits. Nevertheless, to analyze the properties of all reversible (that is, unitary) operations on these qubits, it is helpful to think of the matrices because a lot of useful properties about unitary matrices are known. One important result from this analysis is that every matrix that represents a reversible operation on quantum states can be expressed as a product of the one- and two-qubit gates introduced so far. We say that this set of gates is universal.

For general-purpose computation, it is necessary to have access to arbitrarily many qubits. Instead of assuming that there are infinitely many from the start, it is convenient to have an operation to add a new qubit, namely, **add**. To add a new qubit labeled X in the state $|\mathbf{o}\rangle_X$, apply $\mathbf{add}^{(X)}$ to the current state. This operation can only be used if there is not already a qubit labeled X. To implement the $\mathbf{add}^{(X)}$ operation in the bra-ket notation, we multiply the ket expression for the current state by $|\mathbf{o}\rangle_X$.

Qubit Measurements

In order to classically access information about the state of qubits, we use the measurement operation **meas**. This is an intrinsically probabilistic process that can be applied to any extant qubit. For information processing, one can think of **meas** as a subroutine or function whose output is either 0 or 1. The output is called the measurement outcome. The probabilities of the measurement outcomes are determined by the current state. The state of the qubit being measured is collapsed to the logical state corresponding to the outcome. Suppose we have just one qubit, currently in the state $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. Measurement of this qubit has the effect

$$\text{meas}(\alpha|0\rangle + \beta|1\rangle) = \begin{cases} 0: |0\rangle & \text{with probability } |\alpha|^2 \\ 1: |1\rangle & \text{with probability } |\beta|^2 \end{cases} . \quad (33)$$

The classical output is given before the new state for each possible outcome. This measurement behavior explains why the amplitudes have to define unit length vectors: Up to a phase, they are associated with square roots of probabilities.

For two qubits, the process is more involved. Because of possible correlations between the two qubits, the measurement affects the state of the other one too, similar to conditioning for pbits after one “looks” at one of them. As an example, consider the state

$$|\psi\rangle_{AB} = \frac{2}{3}|01\rangle_{AB} + \frac{i2}{3}|10\rangle_{AB} + \frac{1}{3}|00\rangle_{AB} . \quad (34)$$

To figure out what happens when we measure qubit A, we first rewrite the current state in the form $\alpha|0\rangle_A|\phi_0\rangle_B + \beta|1\rangle_A|\phi_1\rangle_B$, where $|\phi_0\rangle_B$ and $|\phi_1\rangle_B$ are pure states for qubit B. It is always possible to do that. For the example given in Equation (34),

$$\begin{aligned} |\psi\rangle_{AB} &= \frac{2}{3}|0\rangle_A|1\rangle_B + \frac{1}{3}|0\rangle_A|0\rangle_B + \frac{i2}{3}|1\rangle_A|0\rangle_B \\ &= |0\rangle_A \left(\frac{2}{3}|1\rangle_B + \frac{1}{3}|0\rangle_B \right) + |1\rangle_A \frac{i2}{3}|0\rangle_B \\ &= \frac{\sqrt{5}}{3}|0\rangle_A \left(\frac{2}{\sqrt{5}}|1\rangle_B + \frac{1}{\sqrt{5}}|0\rangle_B \right) + \frac{i2}{3}|1\rangle_A|0\rangle_B , \end{aligned} \quad (35)$$

$$\text{so } \alpha = \frac{\sqrt{5}}{3}, \beta = \frac{i2}{3}, |\phi_0\rangle_B = \frac{2}{\sqrt{5}}|1\rangle_B + \frac{1}{\sqrt{5}}|0\rangle_B, \text{ and } |\phi_1\rangle_B = |0\rangle_B .$$

The last step required pulling out the factor of $\sqrt{5/3}$ to make sure that $|\phi_0\rangle_B$ is properly normalized for a pure state. Now, that we have rewritten the state, the effect of measuring qubit A can be given as follows:

$$\text{meas}^{(A)}\left(\alpha|\phi_0\rangle_A|\phi_0\rangle_B + \beta|\phi_1\rangle_A|\phi_1\rangle_B\right) = \begin{cases} 0: |\phi_0\rangle_A|\phi_0\rangle_B & \text{with probability } |\alpha|^2 \\ 1: |\phi_1\rangle_A|\phi_1\rangle_B & \text{with probability } |\beta|^2 \end{cases}. \quad (36)$$

For the example, the measurement outcome is 0 with probability $5/9$, in which case the state collapses to $|\phi_0\rangle_A(1/\sqrt{5}|\phi_0\rangle_B + 2/\sqrt{5}|\phi_1\rangle_B)$. The outcome is 1 with probability $4/9$, in which case the state collapses to $|\phi_1\rangle_A|\phi_1\rangle_B$. The probabilities add up to 1 as they should.

The same procedure works for figuring out the effect of measuring one of any number of qubits. Say we want to measure qubit B among qubits A, B, C, D, currently in state $|\psi\rangle_{ABCD}$. First, rewrite the state in the form $\alpha|\phi_0\rangle_B|\phi_0\rangle_{ACD} + \beta|\phi_1\rangle_B|\phi_1\rangle_{ACD}$, making sure that the ACD superpositions are pure states. Then, the outcome of the measurement is 0 with probability $|\alpha|^2$ and 1 with probability $|\beta|^2$. The collapsed states are $|\phi_0\rangle_B|\phi_0\rangle_{ACD}$ and $|\phi_1\rangle_B|\phi_1\rangle_{ACD}$, respectively.

Probabilities of the measurement outcomes and the new states can be calculated systematically. For example, to compute the probability and state for outcome 0 of $\text{meas}^{(A)}$ given the state $|\psi\rangle_{AB}$, one can first obtain the unnormalized ket expression $|\phi'_0\rangle_B = {}^A\langle\phi_0|\psi\rangle_{AB}$ by using the rules for multiplying kets by bras. The probability is given by $p_0 = {}^B\langle\phi'_0|\phi'_0\rangle_B$, and the collapsed, properly normalized pure state is

$$\frac{|\phi_0\rangle_A|\phi'_0\rangle_B}{\sqrt{p_0}} = \frac{|\phi_0\rangle_A {}^A\langle\phi_0|\psi\rangle_{AB}}{\sqrt{p_0}}. \quad (37)$$

The operator $P_0 = |\phi_0\rangle_A {}^A\langle\phi_0|$ is called a projection operator or projector for short. If we perform the same computation for the outcome 1, we find the projector $P_1 = |\phi_1\rangle_A {}^A\langle\phi_1|$. The two operators satisfy $P_a^2 = P_a$, $P_a^\dagger = P_a$, and $P_0 + P_1 = \mathbb{1}$. In terms of the projectors, the measurement's effect can be written as follows:

$$\text{meas}^{(A)}|\psi\rangle_{AB} = \begin{cases} 0: P_0|\psi\rangle_{AB}/\sqrt{p_0} & \text{with probability } p_0 \\ 1: P_1|\psi\rangle_{AB}/\sqrt{p_1} & \text{with probability } p_1 \end{cases}, \quad (38)$$

where $p_0 = {}^{AB}\langle\psi|P_0|\psi\rangle_{AB}$ and $p_1 = {}^{AB}\langle\psi|P_1|\psi\rangle_{AB}$. In quantum mechanics, any pair of projectors satisfying the properties given above is associated with a potential measurement whose effect can be written in the same form. This is called a binary von Neumann, or projective, measurement.

Mixtures and Density Operators

The measurement operation reads out information from qubits to pbits. What if we discard the pbit that contains the measurement outcome? The result is that the qubits are in a probabilistic mixture of two pure states. Such mixtures are a generalization of pure states. The obvious way to think about a mixture is that we have a probability distribution over pure quantum states. For example, after discarding the pbit and qubit A in Equation (36), we can write the state of B as $\rho = \{|\alpha|^2:|\phi_0\rangle_B, |\beta|^2:|\phi_1\rangle_B\}$, using the notation for probability distributions introduced earlier.

Mixtures frequently form when irreversible operations are used, such as measurement. Except for measurement, the quantum gates we have introduced so far are reversible and therefore transform pure states to pure states so that no mixtures can be formed. One of the fundamental results of reversible classical and quantum computation is that there is no loss in power in using only reversible gates. Specifically, it is possible to change a computation that includes irreversible operations to one that accomplishes the same goal, has only reversible operations, and is efficient in the sense that it uses at most polynomial additional resources. However, the cost of using only reversible operations is not negligible. In particular, for ease of programming and, more important, when performing repetitive error-correction tasks (see the article on this subject on [page 188](#)), the inability to discard or reset qubits can be very inconvenient. We therefore introduce additional operations that enable resetting and discarding.

Although resetting has a so-called thermodynamic cost (think of the heat generated by a computer), it is actually a simple operation. The **reset** operation applied to qubit A can be thought of as the result of first measuring A, then flipping A if the measurement outcome is $|1\rangle$, and finally discarding the measurement result. Using the notation of Equation (36), the effect on a pure state $|\psi\rangle_{AB}$ is given by

$$\text{reset}^{(A)}|\psi\rangle_{AB} = \left\{ |\alpha|^2 : |0\rangle_A |\phi_0\rangle_B, |\beta|^2 : |0\rangle_A |\phi_1\rangle_B \right\} . \quad (39)$$

To apply **reset** to an arbitrary probability distribution, you apply it to each of that distribution's pure states and combine the results to form an expanded probability distribution. The **discard**^(A) operation is **reset**^(A) followed by discarding qubit A. In the expression for the state after **reset**^(A), therefore, all the $|0\rangle_A$ are removed. It is an important fact that every physically realizable quantum operation, whether reversible or not, can be expressed as a combination of **add** operations, gates from the universal set, and **discard** operations.

The representation of mixtures using probability distributions over pure states is redundant. That is, many probability distributions are physically indistinguishable. A nonredundant description of a quantum state can be obtained if density operators are used. The density operator for the mixture ρ in Equation (39) is given by

$$\hat{\rho} = |\alpha|^2 |\phi_0\rangle_B \langle\phi_0| + |\beta|^2 |\phi_1\rangle_B \langle\phi_1| . \quad (40)$$

The general rule for calculating the density operator from a probability distribution is the following: For each pure state $|\phi\rangle$ in the distribution, calculate the operators $|\phi\rangle\langle\phi|$ and sum them weighted by their probabilities.

There is a way to apply gates to the density operators defined by states. If the gate acts by the unitary operator U , then the effect of applying it to $\hat{\rho}$ is given by $U\hat{\rho}U^\dagger$ where U^\dagger is the conjugate transpose of U . (In the bra-ket expression for U , U^\dagger is obtained by replacing all complex numbers by their conjugates, and terms such as $|\phi\rangle\langle\phi|$, by $|\phi\rangle\langle\phi|$.)

The relationship between a qubit's state space and a sphere can be explained in terms of the qubit's density operators. In matrix form, this operator is a 2×2 matrix, which can be written uniquely as a sum $(\mathbb{1} + x\sigma_x + y\sigma_y + z\sigma_z)/2$. One can check that, if the density operator $|\psi\rangle\langle\psi|$ for a qubit's pure state is written as such a sum,

$$|\psi\rangle\langle\psi| = (\mathbb{1} + x\sigma_x + y\sigma_y + z\sigma_z)/2, \quad (41)$$

then the vector (x, y, z) thus obtained is on the surface of the unit sphere in three dimensions. In fact, for every vector (x, y, z) on the unit sphere, there is a unique pure state satisfying Equation (41). Since the density operators for mixtures are arbitrary, convex (that is, probabilistic) sums of pure states, the set of (x, y, z) thus obtained for mixtures fills out the unit ball. The rotations introduced earlier modify the vector (x, y, z) in the expected way, by rotation of the vector around the appropriate axis. See [page 232](#) for more details.

Quantum Computation

The model of computation defined by the one- and two-qubit gates and the operations **add**, **meas**, and **discard** qubits is called the quantum network model. A sequence of instructions for applying these operations is called a quantum network. Quantum computation extends the network model by providing a means for repeating blocks of instructions. Such means can be specified by a formal machine model of computation. There are several such models of classical and quantum computers. One of the best known is the Turing machine, which has a quantum analogue, the quantum Turing machine. This model has its uses for formal studies of computation and complexity but is difficult to program. Fortunately, as mentioned in Part I, there is no loss of computational power if the means for repeating instructions is provided by a classical computer that can apply gates and other operations to qubits. A general quantum algorithm is a program written for such a computer.

There are three practical methods that can be used to write quantum networks and algorithms. The first is to use the names for the different operations and algebraically multiply them. The second is to draw quantum networks, which are pictorial representations of the sequence of steps in time, somewhat like flowcharts without loops. The third is to use a generic programming language enhanced with statements for accessing and modifying quantum bits. The first two methods work well as long as the sequence is short and we do not use many operations that depend on measurement outcomes or require loops. They are often used to describe subroutines of longer algorithms presented either in words or by use of the third method.

To see how to use the different methods and also to illustrate the power of quantum computation, we work out a short quantum algorithm that solves the so-called parity problem.

The Parity Problem. Given is a “black-box” quantum operation $\mathbf{BB}^{(ABC)}$ that has the following effect when applied to a logical basis state:

$$\mathbf{BB}^{(ABC)}|a_A a_B\rangle_{AB}|a_C\rangle_C = |a_A a_B\rangle_{AB}|a_C \oplus (b_A a_A \oplus b_B a_B)\rangle_C, \quad (42)$$

where b_A and b_B are 0 or 1. The actual values of b_A and b_B are unknown. The problem is to determine what b_A and b_B are by using the black box only once.

The terminology and definition of the operation $\mathbf{BB}^{(ABC)}$ require explanation. In computation, we say that an operation is a black box, or an “oracle,” if we have no access whatsoever to how the operation is implemented. In a black-box problem, we are promised that the black box implements an operation from a specified set of operations. In the case of the parity problem, we know that the operation is to add one of four possible parities (see below). The problem is to determine that parity by using the black box in a network. Black-box problems serve many purposes. One is to study the differences between models of computation, just as we are about to do. In fact, black-box problems played a crucial role in the development of quantum algorithms by providing the first and most convincing examples of the power of quantum computers (Bernstein and Vazirani 1993, Simon 1994). Some of these examples involve generalizations of the parity problem. Another purpose of black-box problems is to enable us to focus on what can be learned from the input/output behavior of an operation without having to analyze its implementation. Focusing on the input/output behavior is useful because, in many cases of interest, it is difficult to exploit knowledge of the implementation in order to determine a desirable property of the operation. A classical example is the well-known satisfiability problem, in which we are given a classical circuit with one output bit and we need to determine whether there is an input for which the output is 1. Instead of trying to analyze the circuit, one can look for and use a general-purpose black-box search algorithm to find the satisfying input.

In the definition of the effect of $\mathbf{BB}^{(ABC)}$, the operation \oplus is addition modulo 2, so $1 \oplus 1 = 0$, and all the other sums are as expected. As the state symbols have a numeric meaning now, we will use the number font for states. To see what \mathbf{BB} does, suppose that b_A and b_B are both 1. Then \mathbf{BB} adds (modulo 2) the parity of the logical state in AB to the logical state of C. The parity of a logical state is 0 if the number of 1s is even and 1 if it is odd. The action of \mathbf{BB} for this example is given by

$$\begin{aligned} \mathbf{BB}^{(ABC)}|00\rangle_{AB}|0\rangle_C &= |00\rangle_{AB}|0\rangle_C \\ \mathbf{BB}^{(ABC)}|01\rangle_{AB}|0\rangle_C &= |01\rangle_{AB}|0 \oplus 1\rangle_C \\ &= |01\rangle_{AB}|1\rangle_C \\ \mathbf{BB}^{(ABC)}|10\rangle_{AB}|1\rangle_C &= |10\rangle_{AB}|1 \oplus 1\rangle_C \\ &= |10\rangle_{AB}|0\rangle_C \\ \mathbf{BB}^{(ABC)}|11\rangle_{AB}|0\rangle_C &= |11\rangle_{AB}|0\rangle_C \end{aligned} \quad (43)$$

The action of the black box is extended to superpositions by linear extension. This means that to apply **BB** to a superposition of the logical states, we simply apply it to each logical summand and add the results. Different values of b_A and b_B correspond to different parities. For example, if $b_A = 1$ and $b_B = 0$, then the parity of the state in A is added to the state in C. In this sense, what is added is the parity of a subset of the two qubits AB. Thus, one way of thinking about the problem is that we wish to find out which subset's parity the black box is using.

We can give an algorithm that solves the parity problem using each of the three methods for describing quantum networks. Here is an algebraic description of a solution, **qparity**^(ABC), given as a product of quantum gates that involves one use of the black box. We defer the explanation of why this solution works until after we show how to represent the algorithm pictorially, using quantum networks.

$$\mathbf{qparity}^{(ABC)} = \mathbf{meas}^{(B)} \mathbf{H}^{(B)} \mathbf{meas}^{(A)} \mathbf{H}^{(A)} \mathbf{BB}^{(ABC)} \mathbf{H}^{(C)} \mathbf{not}^{(C)} \mathbf{add}^{(C)} \mathbf{H}^{(B)} \mathbf{add}^{(B)} \mathbf{H}^{(A)} \mathbf{add}^{(A)} . \quad (44)$$

The output of the algorithm is given by the classical outputs of the measurements of qubit A, which yield b_A , and of qubit B, which yield b_B . As is conventional, in writing products of linear operators, the order of application in Equation (44) is right to left, as in a product of matrices applied to a column vector. This order of terms in a product is, however, counterintuitive, particularly for operations to be performed sequentially. It is therefore convenient to use left to right notation, as is done in describing laser or radio-frequency pulse sequences, and to put dots between gates to indicate left to right order:

$$\mathbf{qparity}^{(ABC)} = \mathbf{add}^{(A)} . \mathbf{H}^{(A)} . \mathbf{add}^{(B)} . \mathbf{H}^{(B)} . \mathbf{add}^{(C)} . \mathbf{not}^{(C)} . \mathbf{H}^{(C)} . \mathbf{BB}^{(ABC)} . \mathbf{H}^{(A)} . \mathbf{meas}^{(A)} . \mathbf{H}^{(B)} . \mathbf{meas}^{(B)} . \quad (45)$$

In this representation, the first operation is **add**^(A), the second is **H**^(A) (the Hadamard gate on qubit A), and so on.

The algebraic specification of the algorithm as products of gates does not make it easy to see why the algorithm works. It is also difficult to see which operations depend on each other. Such dependencies are used to determine whether the operations can be parallelized. Quantum networks make these tasks simpler. The quantum network for the above sequence is shown in Figure 2.

To understand how the quantum network illustrated in Figure 2 solves the parity problem, we can follow the states as the network is executed from left to right, using the indicated checkpoints. Using vector notation for the states, at checkpoint 1, the state is

$$|\psi\rangle_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad (46)$$

where we used Kronecker product notation to denote the states of A, B, and C in this order. In the next time step, the network involves applying Hadamard gates—see Equation (13)—to A and B and a **not** gate—see Equation (9)—to C. At checkpoint 2, this operation results in the state

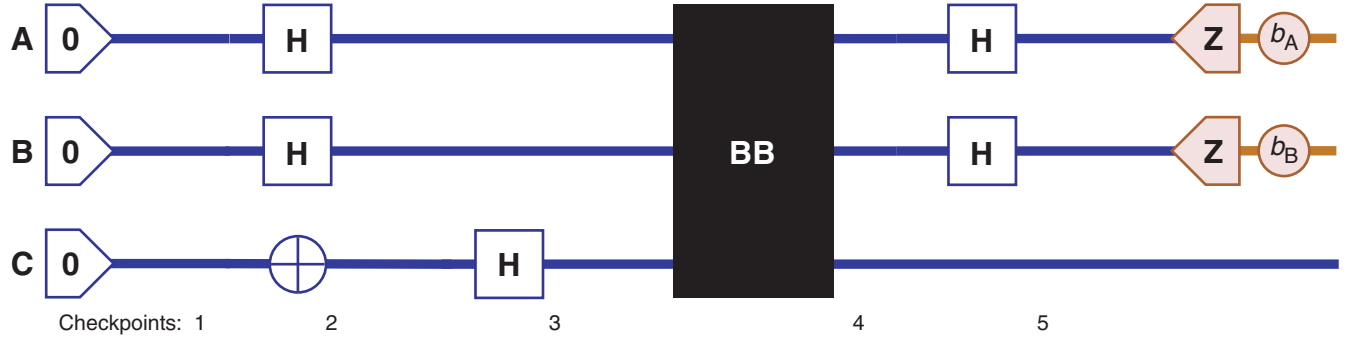


Figure 2. Quantum Network for Solving the Parity Problem

A quantum network has a line (horizontal in this case) for each qubit. The line can be thought of as the timeline for the qubit and is shown in blue. Each gate is drawn as a box, circle, or other element intercepting the lines of the qubits it acts on. In this case, time runs from left to right. Each qubit's timeline starts at the point where it is added. In this example, the qubits' timelines end when they are measured, at which point a classical bit (brown timeline) containing the measurement outcome is introduced. The operation BB is illustrated as a black box. The numbers underneath the network refer to checkpoints used to explain how the network solves the parity problem.

$$|\psi\rangle_2 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \otimes \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \otimes \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (47)$$

Next, a Hadamard gate is applied to C, so that at checkpoint 3, we have

$$|\psi\rangle_3 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \otimes \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} \otimes \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}. \quad (48)$$

The next event involves applying the black box. To understand what happens, note that the effect of the black box can be described as, “apply **not** to C if the parity according to b_A and b_B of the logical state of AB is 1.” The current state of C is such that, if **not** is applied, only the sign changes:

$$\begin{aligned} \text{not} \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} \\ &= -\begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}. \end{aligned} \quad (49)$$

Now, AB is in a superposition of each of the logical states, and conditional on the logical state and the (hidden) parity, the sign changes. As a result, although the state of C does not change, a phase is “kicked back” to AB. A generalization of this effect is at the heart of Alexei Kitaev’s version of Peter Shor’s quantum factoring algorithm (see the article “From Factoring to Phase Estimation” on [page 38](#)). At the next checkpoint, and after some arithmetic to check which logical states change sign, we can write the state as

$$|\psi\rangle_4 = \begin{pmatrix} 1/\sqrt{2} \\ (-1)^{b_A}/\sqrt{2} \end{pmatrix} \otimes \begin{pmatrix} 1/\sqrt{2} \\ (-1)^{b_B}/\sqrt{2} \end{pmatrix} \otimes \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}. \quad (50)$$

Notice that qubits A and B are in orthogonal states for different values of b_A and b_B . It suffices to apply the Hadamard transform again to A and B to get

$$|\psi\rangle_4 = \begin{pmatrix} 1-b_A \\ b_A \end{pmatrix} \otimes \begin{pmatrix} 1-b_B \\ b_B \end{pmatrix} \otimes \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}. \quad (51)$$

Measurements of A and B now reveal the previously unknown b_A and b_B .

As can be seen, the visual representation of a quantum network eases the tasks of following what happens. This is why it is used extensively for presenting basic subroutines and algorithms in quantum computation. A guide to the commonly used network elements is given in Table I.

When designing or describing complicated algorithms for quantum computers, providing everything in terms of quantum networks can become difficult, particularly when an important part of the algorithm consists of computations that are best done on a classical computer. For example, a full description of Shor's algorithm for factoring integers (see the article "From Factoring to Phase Estimation" on [page 38](#)) includes a significant amount of classical preprocessing, which determines choices made in the quantum algorithm, and classical postprocessing, which computes a factor from the measured result by a continued fraction algorithm. For such algorithms, one can use a programming language similar to Pascal, BASIC, or C enhanced with statements to access quantum bits and to apply quantum operations. For algorithm design, computer scientists often use a semiformal language called pseudocode (Cormen et al. 1990). With a simple extension called quantum pseudocode, the algorithm for the parity problem can be written as follows:

BBPARITY(BB)

Input: Access to a quantum black box **BB** that acts on three qubits by adding a parity function of the first two qubits to the third

Output: The two bits b_A and b_B of the parity function

```

foreach  $i \in \{A, B, C\}$ 
   $\lceil a_i \rceil \leftarrow |0\rangle$ 
  C: Initialize three one-qubit registers  $\lceil a_i \rceil$ ,  $i = A, B, C$ .

  The corner bracket annotation declares  $a_i$  as a quantum register.

  end

   $\lceil a_C \rceil \leftarrow \sigma_x \lceil a_C \rceil$ 

  foreach  $i \in \{A, B, C\}$ 
     $\lceil a_i \rceil \leftarrow \mathbf{H} \lceil a_i \rceil$ 
  end

   $\lceil a \rceil \leftarrow \mathbf{BB} \lceil a \rceil$ 

  C:  $\lceil a \rceil$  refers to the three-qubit register consisting of the  $\lceil a_i \rceil$ .

  foreach  $i \in \{A, B\}$ 
     $\lceil a_i \rceil \leftarrow \mathbf{H} \lceil a_i \rceil$ 
```

```

 $b_i \leftarrow \text{meas } \lceil a_i \rceil$ 

end

return  $b_A, b_B$ 

end

```

Any classical programming language can be extended with statements to access and manipulate quantum registers.

Now, that we have looked at the quantum solution to the parity problem, let us consider the question of the least number of black-box applications required by a classical algorithm: Each classical use of the black box can only give us one bit of information. In particular, one use of the black box with input $a_A a_B$ reveals only the parity of $a_A a_B$ according to the hidden parameters b_A and b_B . Each use of the black box can therefore only help us distinguish between two subsets of the four possible parities. At least two uses of the black box are therefore necessary. Two uses are also sufficient. To determine which of the four parities is involved, use the black box first with input $a_A a_B = 10$ and then with input $a_A a_B = 01$. As a result of this argument, one can consider the parity problem as a simple example of a case in which there is a more efficient quantum algorithm than is possible

Table I. Quantum Network Elements










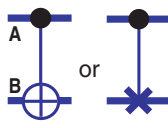

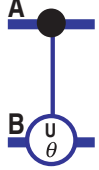

Gate Names and Their Abbreviations	Gate Symbols	Algebraic Form	Matrix Form
Add/Prepare, add		If applied to an existing qubit $\{ 0\rangle\langle 0 , 0\rangle\langle 1 \}$ (operator mixture)	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ or $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$
Measure, meas		$\{0: 0\rangle\langle 0 , 1: 1\rangle\langle 1 \}$	$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ or $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$
Not, not , σ_x	 or 	$ 0\rangle\langle 1 + 1\rangle\langle 0 $	$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
Hadamard, H		$e^{-i\sigma_y\pi/4}\sigma_z$	$\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$
Phase Change, S ($e^{i\phi}$)		$e^{i\phi/2}e^{-i\sigma_z\phi/2}$	$\begin{pmatrix} 1 & 0 \\ 0 & e^{i\phi} \end{pmatrix}$

Table I. (Continued)

Gate Names and Their Abbreviations	Gate Symbols	Algebraic Form	Matrix Form
z-Rotation, \mathbf{Z}_ϕ		$e^{-i\sigma_z\phi/2}$	$\begin{pmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{pmatrix}$
y-Rotation, \mathbf{Y}_θ		$e^{-i\sigma_y\theta/2}$	$\begin{pmatrix} \cos(\theta/2) & -\sin(\theta/2) \\ \sin(\theta/2) & \cos(\theta/2) \end{pmatrix}$
x-Rotation, \mathbf{X}_θ		$e^{-i\sigma_x\theta/2}$	$\begin{pmatrix} \cos(\theta/2) & -i\sin(\theta/2) \\ -i\sin(\theta/2) & \cos(\theta/2) \end{pmatrix}$
Controlled not, cnot		$ 0\rangle_A \langle 0 + 1\rangle_A \langle 1 \sigma_x^{(B)}$ $e^{-i\sigma_z^{(A)}\pi/4} e^{-i/2(1-\sigma_z^{(A)})\sigma_x^{(B)}\pi/2}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$
zz-Rotation, $(\mathbf{ZZ})_\theta$		$e^{-i\sigma_z^{(A)}\sigma_z^{(B)}\theta/2}$	$\begin{pmatrix} e^{-i\theta/2} & 0 & 0 & 0 \\ 0 & e^{i\theta/2} & 0 & 0 \\ 0 & 0 & e^{i\theta/2} & 0 \\ 0 & 0 & 0 & e^{-i\theta/2} \end{pmatrix}$
Controlled Rotation, cU $_\theta$		$ 0\rangle_A \langle 0 + 1\rangle_A \langle 1 e^{-i\sigma_U^{(B)}\theta/2}$	$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & & e^{-i\sigma_U\theta/2} \\ 0 & 0 & & \end{pmatrix}$
Toffoli Gate, c²not		$\mathbb{1} - 11\rangle_{AB} \langle 11 + 11\rangle_{AB} \langle 11 \sigma_x^{(C)}$	

classically. However, it is worth noting that the comparison is not entirely fair: A truly classical oracle answering parity questions or implementing the black box on the states of classical bits is useless to a quantum algorithm. To take advantage of such an algorithm, it must be possible to use superpositions that are not implicitly collapsed. Collapse can happen if the oracle makes a measurement or otherwise “remembers” the question that it was asked.

Resource Accounting

When trying to solve a problem using quantum information processing, an important issue is to determine what physical resources are available and how much of each resource is needed for the solution. As mentioned before, in classical information, the primary resources are bits and operations. The number of bits used by an algorithm is its space requirement; the number of operations used, its time requirement. If parallel computation is available, one can distinguish between the total number of operations (work) and the number of parallel steps (time).

When quantum information processing is used, the classical resources are still relevant for running the computer that controls the quantum system and performs any preprocessing and postprocessing tasks. The main quantum resources are analogous to the classical ones: Quantum space is the number of qubits needed, and quantum time, the number of quantum gates. Because it turns out that reset operations have a thermodynamic cost, one can count irreversible quantum operations separately. This accounting of the resource requirements of algorithms and of the minimum resources needed to solve problems forms the foundation of quantum complexity theory.

As a simple example of resource accounting, consider the algorithm for the parity problem. No classical computation is required to decide which quantum gates to apply or to determine the answer from the measurement. The quantum network consists of a total of 11 quantum gates (including **add** and **meas** operations) and one oracle call (the application of the black box). In the case of oracle problems, one usually counts the number of oracle calls first, as we have done in discussing the algorithm. The network is readily parallelized to reduce the time resource to 6 steps.

Part III: Advantages of Quantum Information

The notion of quantum information as explained in this primer was established in the 1990s. It emerged from research focused on understanding how physics affects our capabilities to communicate and process information. The recognition that usable types of information need to be physically realizable was repeatedly emphasized by Rolf Landauer, who proclaimed that “information is physical” (1991). Beginning in the 1960s, Landauer studied the thermodynamic cost of irreversible operations in computation (1961). Charles Bennett showed that, by using reversible computation, this cost can be avoided (1973). Limitations of measurement in quantum mechanics were investigated early by researchers such as John von Neumann (1932a and 1932b) and later by Alexander Holevo (1973b) and Carl Helstrom (1976). Holevo introduced the idea of quantum communication channels and found bounds on their capacity for transmitting classical information (1973a). Initially, most work focused on determining the physical limitations placed on classical information processing. The fact that pairs of two-level systems can have correlations not possible for classical systems was proved by John Bell (1964). Subsequently, indications that quantum mechanics offers advantages to information processing came from Stephen Wiesner’s studies of cryptographic applications in the late 1960s. Wiesner’s work was not recognized, however, until the 1980s,

when Bennett, Gilles Brassard, Seth Breidbart, and Wiesner introduced (1982) the idea of quantum cryptography, which can be used to communicate in secret.

Initially, the term quantum computation was mostly used to refer to classical computers realized with quantum mechanical systems. In the 1980s, Paul Benioff (1980), Richard Feynman (1982), and Yuri Manin (1980) introduced the idea of a quantum computer based on quantum information. They noted that the apparent exponential complexity of simulating quantum mechanics on a classical computer might be overcome if one could use a computer based on quantum mechanics. A formal model of quantum Turing machines was soon defined by David Deutsch (1985), who later also introduced quantum networks (1989). Deutsch and Richard Jozsa (1992) were the first to introduce a black-box problem that could be solved deterministically on a quantum computer in fewer steps than on a classical computer.

In spite of suggestions that it could lead to large efficiency improvements in simulating physics, quantum information processing was still largely an academic subject. Based on work by Ethan Bernstein and Umesh Vazirani (1993) that formalized quantum complexity theory, Dan Simon (1994) showed that, for black-box problems, quantum computers can be exponentially more efficient than classical deterministic or probabilistic computers, giving the first indication of a strong advantage for quantum information processing. It was Shor's algorithm for factoring large integers (1994 and 1997) that finally convinced a larger community that quantum information was more than just a tool for realizing classical computers. This change in attitude was in no small part due to the fact that the security of commonly used cryptographic protocols is based on the difficulty of factoring.

At that point, it was still generally believed that the fragility of quantum states made it unlikely for reasonably large quantum computers to be realized in practice. But the discovery by Shor (1995) and Andrew Steane (1996) that quantum error correction was possible soon changed that view (for an introductory overview, see the article on quantum error correction on page 188).

Because the usefulness and realizability of quantum information has been recognized, the science of quantum information processing is a rapidly growing field. As quantum information becomes increasingly accessible by technology, its usefulness will be more apparent. The next few sections discuss what we currently know about applications of quantum information processing. Refer to Michael Nielsen and Isaac Chuang (2001) as a useful reference text on quantum computation and information with historical notes.

Quantum Algorithms

Shor's factoring algorithm, which precipitated much of the current work in quantum information processing, is based on a quantum realization of the fast Fourier transform. The most powerful version of this technique is now represented by the phase estimation algorithm of Kitaev (1995) as formalized by Richard Cleve et al. (1998). (For an explanation, see the article "From Factoring to Phase Estimation" on page 38.) The best-known application of quantum factoring is cryptanalysis, where it allows efficiently breaking the currently used public-key cryptographic codes. Whether there are any constructive applications of quantum factoring and its generalizations remains to be determined. For users of public-key cryptography, a crucial question is, "How long can public-key codes based on factoring continue to be used safely?" To attempt an answer to this question, one can note that to break a code with a typical key size of 1000 bits requires more than 3000 qubits and 10^8 quantum gates, which is well out of reach of current technology. However, it is conceivable that a recording of encrypted information transmitted in 2000 can be broken in the next "few" decades.

Shor's quantum factoring algorithm was not the first with a significant advantage over classical algorithms. The first proposed quantum algorithms with this property were for simulating quantum mechanical systems. These algorithms simulate the evolution of a reasonably large number of interacting quantum particles—for example, the electrons and nuclei in a molecule. The algorithms' outputs are what would be measurable physical quantities of the system being simulated. The known methods for obtaining these quantities on classical computers scale exponentially with the number of particles, except in special cases.

The idea of using quantum computers for simulating quantum physics spurred the work that eventually led to the quantum factoring algorithm. However, that idea did not have the broad scientific impact that the quantum factoring algorithm had. One reason is that, because of its cryptographic applications, factoring is a heavily studied problem in theoretical computer science and cryptography. Because so many people have tried to design efficient algorithms for factoring and failed, the general belief that factoring is hard for classical computers has a lot of credibility. In contrast, a quantum physics simulation has no simple formulation as an algorithmic problem suitable for study in theoretical computer science. Furthermore, many researchers still believe that the physically relevant questions can be answered with efficient classical algorithms, requiring only more cleverness on the part of algorithm designers. Another reason for the lack of impact is that many of the fundamental physical quantities of interest are not known to be efficiently accessible even on quantum computers. For example, one of the first questions about a physical system with a given Hamiltonian (energy observable) is, "What is the ground-state energy?" It is known that the ability to efficiently answer this question for physically reasonable Hamiltonians leads to efficient algorithms for hard problems, such as the traveling salesman or the scheduling problems. In spite of occasional claims to the contrary, an efficient quantum solution to these problems is widely considered unlikely.

Most quantum algorithms for physics simulations are based on a direct emulation of a quantum mechanical system's evolution. The focus of the original proposals by Feynman and others was on how to implement the emulation using a suitable formulation of general-purpose quantum computers. After such computers were formalized by Deutsch, the implementation of the emulation was generalized and refined by Seth Lloyd (1996), Wiesner (1996), and Christof Zalka (1998). The ability to emulate the evolution of quantum systems is actually widely used by classical Monte Carlo algorithms for simulating physics. In those algorithms, state amplitudes are, in effect, represented by expectations of random variables that are computed during the simulation. As in the case of quantum algorithms for physics emulation, Monte Carlo algorithms efficiently evolve the representation of the quantum system. The inefficiency of the classical algorithm arises only in determining a physical quantity of interest. In the case of Monte Carlo algorithms, the measurement of a physical quantity suffers from the so-called sign problem, often resulting in exponentially large, random errors that can be reduced only by repeating the computation exponentially many times. In contrast, the quantum algorithms for emulation can determine many (but not all) of the interesting physical quantities with polynomially bounded statistical errors. How to efficiently implement measurements of these quantities has been the topic of more recent work in this area, much of which is based on variants of the phase-estimation algorithm (Terhal and DiVincenzo 2000, Knill and Laflamme 1998, Abrams and Lloyd 1999, Ortiz et al. 2001, Miquel et al. 2002).

Although several researchers have suggested that there are interesting quantum physics simulations that can be implemented with well below 100 qubits, one of the interesting problems in this area of research is to come up with a specific simulation algorithm using small numbers of qubits and quantum gates, an algorithm that computes an interesting physical quantity not easily obtainable using available classical computers.

Another notable algorithm for quantum computers, unstructured quantum search, was described by Lov Grover (1996). Given is a black box that computes a binary function f on inputs x with $0 \leq x < N$. The function f has the property that there is a unique input a for which $f(a) = 1$. The standard quantum version of this black box implements the transformation $\hat{f}|x\rangle|b\rangle = |x\rangle|b \oplus f(x)\rangle$, where b is a bit and $b \oplus f(x)$ is computed modulo 2. Unstructured quantum search finds a quadratically faster, that is, in time of order $N^{1/2}$, than the best classical black-box search, which requires time of order N . The context for this algorithm is the famous $P \neq NP$ conjecture, which is captured by the following algorithmic problem: Given is a classical circuit C that computes an output. Is there an input to the circuit for which the circuit's output is 1? Such an input is called a satisfying input or assignment. For any given input, it is easy to check the output, but an efficient algorithm that finds a satisfying input is conjectured to be impossible. This is the $P \neq NP$ conjecture. Generalizations of Grover's search algorithm—amplitude amplification (Brassard et al. 1998)—allow finding satisfying inputs faster than naive, classical search does, which tries every possible input in some, possibly random, order. It is worth noting, however, that if sufficient classical parallelism is available, quantum search loses many of its advantages.

The three algorithms just described capture essentially all the known algorithmic advantages of quantum computers. Almost all algorithms that have been described are applications of phase estimation or of amplitude amplification. These algorithms well justify developing special-purpose quantum information-processing technology. Will general-purpose quantum computers be useful? More specifically, what other algorithmic advantages do quantum computers have?

Quantum Communication

Quantum communication is an area in which quantum information has proven (rather than conjectured) advantages. The best-known application is quantum cryptography, whereby two parties, Alice and Bob, can generate a secret key using a quantum communication channel (for example, photons transmitted in optical fiber) and an authenticated classical channel (for example, a telephone line). Any attempt at learning the key by eavesdropping is detected. A quantum protocol for generating a secret key is called a quantum-key-exchange protocol. There are no equally secure means for generating a secret key by using only classical deterministic channels. Few quantum operations are needed to implement quantum key exchange, and as a result, there are working prototype systems (Hughes et al. 2000, Townsend 1998, Ribordy et al. 2001). To overcome the distance limitations (tens of kilometers) of current technology requires the use of quantum error correction and hence more demanding quantum technology.

Quantum key exchange is one of an increasing number of multiparty problems that can be solved more efficiently with quantum information. The area of research concerned with how several parties at different locations can solve problems while minimizing communication resources is called communication complexity. For quantum communication complexity (Cleve and Burhman 1997), the communication resources include either shared entangled qubits or a means for transmitting quantum bits. A seminal paper by Howard Burhman, Cleve, and Wim van Dam (2000) shows how the nonclassical correlations present in maximally entangled states lead to protocols based on such states that are more efficient than any classical deterministic or probabilistic protocol achieving the same goal. Ran Raz (1999) showed that there is an exponential improvement in communication resources for a problem in which Alice and Bob have to answer a question about the relationship between a vector known to Alice and a matrix known to Bob. Although this problem is artificial, it suggests that there are potentially useful advantages to be gained from quantum information in this setting.

Quantum Control

According to Moore's law of semiconductor technology, the size of transistors is decreasing exponentially, by a factor of about .8 every year. If this trend continues, then over the next few decades, devices will inevitably be built whose behavior will be primarily quantum mechanical. For the purpose of classical computation, the goal is to remove the quantum behavior and stabilize classical information. But quantum information offers an alternative: It is possible to directly use quantum effects to advantage. Whether or not this alternative is useful (and we believe it is), the ideas of quantum information can be used to systematically understand and control quantum mechanical systems.

The decreasing size of semiconductor components is a strong motivation to strive for better understanding the behavior of condensed-matter quantum mechanical systems. But there is no reason to wait for Moore's law: There are a rapidly increasing number of experimental systems in which quantum mechanical effects are being used and investigated. Examples include many optical devices (lasers, microwave cavities, entangled photon pairs), nuclear magnetic resonance with molecules or in solid state, trapped ion or atom systems, Rydberg atoms, superconducting devices (Josephson junctions and SQUIDs), and spintronics (electron spins in semiconductor devices). Many of these systems are being considered as candidates for realizing quantum information processing. Yet, regardless of the future of quantum information processing, there is ample motivation for studying these systems.

Outlook

The science of quantum information processing is promising a significant impact on how we process information, solve algorithmic problems, engineer nanoscale devices, and model fundamental physics. It is already changing the way we understand and control matter at the atomic scale, making the quantum world more familiar, accessible, and understandable. Whether or not we do most of our everyday computations by using the classical model, it is likely that the physical devices that support these computations will exploit quantum mechanics and integrate the ideas and tools that have been developed for quantum information processing. ■

Acknowledgment

We thank Nikki Cooper and Ileana Buican for their extensive encouragement and editorial help.

Further Reading

- Abrams, D. S., and S. Lloyd. 1999. Quantum Algorithm Providing an Exponential Speed Increase for Finding Eigenvalues and Eigenvectors. *Phys. Rev. Lett.* **83**: 5162.
- Barenco, A., C. H. Bennett, R. Cleve, D. P. DiVincenzo, N. Margolus, P. Shor, et al. 1995. Elementary Gates for Quantum Computation. *Phys. Rev. A* **52**: 3457.
- Bell, J. S. 1964. On the Einstein-Podolsky-Rosen Paradox. *Phys.* **1**: 195.
- Benioff, P. 1980. The Computer as a Physical System: A Microscopic Quantum Mechanical Hamiltonian Model of Computers as Represented by Turing Machines. *J. Stat. Phys.* **22**: 563.
- Bennett, C. H. 1973. Logical Reversibility of Computation. *IBM J. Res. Dev.* **17**: 525.
- Bennett, C. H., G. Brassard, S. Breidbart, and S. Wiesner. 1982. Quantum Cryptography, or Unforgeable Subway Tokens. In *Advances in Cryptology: Proceedings of Crypto '82*. Edited by D. Chaun, R. L. Rivest, and A. T. Sherman, p. 267. New York: Plenum Press.



Emanuel (Manny) Knill received his Ph. D. in pure mathematics from the University of Colorado at Boulder in 1991. Since 1992, he has been with Los Alamos National Laboratory. Manny has worked on various aspects of quantum information processing since 1995. (Photo of Augustus)

Contact Information

E. Knill: knill@lanl.gov

R. Laflamme: laflamme@iqc.ca

H. Barnum: barnum@lanl.gov

D. Dalvit: dalvit@lanl.gov

J. Dziarmaga: ufjacekd@th.if.uj.edu.pl

J. Gubernatis: jg@lanl.gov

L. Gurvits: gurvits@lanl.gov

G. Ortiz: g_ortiz@lanl.gov

L. Viola: lviola@lanl.gov

W. Zurek: whz@lanl.gov

- Bernstein, E., and U. Vazirani. 1993. Quantum Complexity Theory. In *Proceedings of the 25th Annual ACM Symposium on the Theory of Computing 1993*, p. 11. New York: ACM Press.
- Bolker, E. D. 1970. *Elementary Number Theory: An Algebraic Approach*. New York: W. A. Benjamin, Inc.
- Brassard, G., P. Hoyer, and A. Tapp. 1998. Quantum Counting. In *Automata, Languages and Programming, Proceedings of ICALP '98*, Vol. 1443 of *Lecture Notes in Computer Science*. Edited by K. G. Larsen, S. Skyum, and G. Winskel, p. 820. Berlin: Springer Verlag.
- Buhrman, H., R. Cleve, and W. Van Dam. 2000. Quantum Entanglement and Communication Complexity. *SIAM J. Comput.* **30**: 1829.
- Cleve, R., and H. Buhrman. 1997. Substituting Quantum Entanglement for Communication. *Phys. Rev. A* **56**: 1201.
- Cleve, R., A. Ekert, C. Macchiavello, and M. Mosca. 1998. Quantum Algorithms Revisited. *Proc. R. Soc. London, Ser. A* **454**: 339.
- Cormen, T. H., C. B. Leiserson, and R. L. Rivest. 1990. *Introduction to Algorithms*. Cambridge, MA: MIT Press.
- Deutsch, D. 1985. Quantum Theory, The Church-Turing Principle and the Universal Quantum Computer. *Proc. R. Soc. London, Ser. A* **400**: 97.
- . 1989. Quantum Computational Networks. *Proc. R. Soc. London, Ser. A* **425**: 73.
- Deutsch, D., and R. Jozsa. 1992. Rapid Solution of Problems by Quantum Computation. *Proc. R. Soc. London, Ser. A* **439**: 553.
- Ekert, A. 1998. From Quantum Code-Making to Quantum Code-Breaking. In *The Geometric Universe*, p. 195. Oxford: Oxford University Press.
- Feynman, R. P. 1982. Simulating Physics with Computers. *Int. J. Theor. Phys.* **21**: 467.
- Griffiths, R. B., and C.-S. Niu. 1996. Semiclassical Fourier Transform for Quantum Computation. *Phys. Rev. Lett.* **76**: 3228.
- Grover, L. K. 1996. A Fast Quantum Mechanical Algorithm for Database Search. In *Proceedings of the 28th Annual ACM Symposium on the Theory of Computation*, p. 212, New York: ACM Press.
- Gupta, R., S. A. Smolka, and S. Bhaskar. 1994. On Randomization in Sequential and Distributed Algorithms. *ACM Comput. Surveys* **26**: 7.
- Hardy, G. H., and E. M. Wright. 1979. *An Introduction to the Theory of Numbers*. Fifth edition. London: Oxford University Press.
- Helstrom, C. W. 1976. *Quantum Detection and Estimation Theory. Mathematics in Science and Engineering* Vol. 123. New York: Academic Press.
- Holevo, A. S. 1973a. Bounds for the Quantity of Information Transmitted by a Quantum Communication Channel. *Probl. Inf. Transm.* **9**: 177.
- . 1973b. Statistical Problems in Quantum Physics. In *Proceedings of the Second Japan-USSR Symposium on Probability Theory, Lecture Notes in Mathematics* Vol. 330. Edited by G. Maruyama and Y. V. Prokhorov, p. 104. Berlin: Springer Verlag.
- Hughes, R. J., G. L. Morgan, and C. G. Peterson. 2000. Quantum Key Distribution Over a 48-km Optical Fibre Network. *J. Modern Opt.* **47**: 533.
- Kitaev, A. Yu. 1995. Quantum Measurements and the Abelian Stabilizer Problem. [Online]: [http://eprints.lanl.gov. \(quant-ph/9511026\)](http://eprints.lanl.gov. (quant-ph/9511026)).
- Knill, E., and R. Laflamme. 1998. On the Power of One Bit of Quantum Information. *Phys. Rev. Lett.* **81**: 5672.
- Landauer, R. 1961. Irreversibility and Heat Generation in the Computing Process. *IBM J. Res. Dev.* **5**: 183.
- . 1991. Information is Physical. *Phys. Today* **44**: 22.
- Lloyd, S. 1996. Universal Quantum Simulators. *Science* **273**: 1073.
- Manin, Y. I. 1980. *The Computable and the Not Computable*. Moscow: Sovetskoye Radio. (In Russian).
- Miquel, C., J. P. Paz, M. Saraceno, E. Knill, R. Laflamme, and C. Negrevergne. 2002. Interpretation of Tomography and Spectroscopy as Dual Forms of Quantum Computations. *Nature* **418**: 59.
- Nielsen, M. A., and I. L. Chuang. 2001. *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.

- Ortiz, O., J. E. Gubernatis, E. Knill, and R. Laflamme. 2001. Quantum Algorithms for Fermionic Simulations. *Phys. Rev. A* **64**: 022319.
- Papadimitriou, C. H. 1994. *Computational Complexity*. Reading, MA: Addison-Wesley.
- Raz, R. 1999. Exponential Separation of Quantum and Classical Communication Complexity. In *Proceedings of the 31st Annual ACM Symposium on the Theory of Computation (STOC)*, p. 358. El Paso, TX: ACM Press.
- Ribordy, O., J. Brendel, J.-D. Gautier, N. Gisin, and H. Zbinden. 2001. Long-Distance Entanglement-Based Quantum Key Distribution. *Phys. Rev. A* **63**: 012309.
- Shor, P. W. 1994. Algorithms for Quantum Computation: Discrete Logarithms and Factoring. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, p. 124. Los Alamitos, CA: IEEE Press.
- . 1995. Scheme for Reducing Decoherence in Quantum Computer Memory. *Phys. Rev. A* **52**: 2493.
- . 1997. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer. *SIAM J. Comput.* **26**: 1484.
- Simon, D. R. 1994. On the Power of Quantum Computation. In *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, p. 116. Los Alamitos, CA: IEEE Press.
- Steane, A. 1996. Multiple Particle Interference and Quantum Error Correction. *Proc. R. Soc. London, Ser. A* **452**: 2551.
- Terhal, B. M., and D. P. DiVincenzo. 2000. Problem of Equilibration and the Computation of Correlation Functions on a Quantum Computer. *Phys. Rev. A* **61**: 022301.
- Townsend, P. D. 1998. Quantum Cryptography on Optical Fiber Networks. *Opt. Fiber Tech.: Mat., Dev., Sys.* **4**: 345.
- von Neumann, J. 1932a. Der Messprozess. Ch. VI. In *Mathematische Grundlagen der Quantenmechanik*. Berlin: Springer Verlag.
- . 1932b. "Messung und Reversibilität." Allgemeine Betrachtungen. Ch. V. In *Mathematische Grundlagen der Quantenmechanik*. Berlin: Springer Verlag.
- Wiesner, S. 1983. Conjugate Coding. *Sigact News* **15**: 78.
- . 1996. Simulations of Many-Body Quantum Systems by a Quantum Computer. [Online]: [http://eprints.lanl.gov. \(quant-ph/9603028\)](http://eprints.lanl.gov. (quant-ph/9603028)).
- Yao, A. 1993. Quantum Circuit Complexity. In *Proceedings of the 34th Annual Symposium on Foundations of Computer Science*, p. 352. Los Alamitos, CA: IEEE Press.
- Zalka, C. 1998. Simulating Quantum Systems on a Quantum Computer. *Proc. R. Soc. London, Ser. A* **454**: 313.

Glossary

- Algorithm.** A set of instructions to be executed by a computing device. What instructions are available depends on the computing device. Typically, instructions include commands for manipulating the contents of memory and means for repeating blocks of instructions indefinitely or until a desired condition is met.
- Amplitude.** A quantum system with a chosen orthonormal basis of "logical" states $|i\rangle$ can be in any superposition $\sum_i \alpha_i |i\rangle$ of these states, where $\sum_i |\alpha_i|^2 = 1$. In such a superposition, the complex numbers α_i are called the amplitudes. Note that the amplitudes depend on the chosen basis.
- Ancillas.** Helper systems used to assist in a computation involving other information systems.
- Bell basis.** For two qubits A and B, the Bell basis consists of the four states $1/\sqrt{2}(|00\rangle_{AB} \pm |11\rangle_{AB})$ and $1/\sqrt{2}(|01\rangle_{AB} \pm |10\rangle_{AB})$.
- Bell states.** The members of the Bell basis.
- Bit.** The basic unit of deterministic information. It is a system that can be in one of two possible states, 0 and 1.

- Bit sequence.** A way of combining bits into a larger system whose constituent bits are in a specific order.
- Bit string.** A sequence of 0s and 1s that represents a state of a bit sequence. Bit strings are the words of a binary alphabet.
- Black box.** A computational operation whose implementation is unknown. Typically, a black box implements one of a restricted set of operations, and the goal is to determine which of these operations it implements by using it with different inputs. Each use of the black box is called a “query.” The smallest number of queries required to determine the operation is called the “query complexity” of the restricted set. Determining the query complexity of sets of operations is an important area of computational complexity.
- Bloch sphere.** The set of pure states of a qubit represented as points on the surface of the unit sphere in three dimensions.
- Bra.** A state expression of the form $\langle\psi|$ considered to be the conjugate transpose of the ket expression $|\psi\rangle$.
- Bra-ket notation.** A way of denoting states and operators of quantum systems with kets (for example, $|\psi\rangle$) and bras (for example, $\langle\phi|$).
- Circuit.** A combination of gates applied to information units in a prescribed order.
To draw circuits, one often uses a convention for connecting and depicting gates.
See also “network.”
- Circuit complexity.** The circuit complexity of an operation on a fixed number of information units is the smallest number of gates required to implement the operation.
- Classical information.** The type of information based on bits and bit strings and more generally on words formed from finite alphabets. This is the information used for communication between people. Classical information can refer to deterministic or probabilistic information, depending on the context.
- Computation.** The execution of the instructions provided by an algorithm.
- Computational states.** See “logical states.”
- Computer.** A device that processes information.
- Density matrix or operator.** A representation of pure and mixed states without redundancy. For a pure state $|\psi\rangle$, the corresponding density operator is $|\psi\rangle\langle\psi|$.
A general density operator is a probabilistic combination $\sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$, with $\sum_i \lambda_i = 1$.
- Deterministic information.** The type of information that is based on bits and bit strings.
Deterministic information is classical, but it explicitly excludes probabilistic information.
- Distinguishable states.** In quantum mechanics, two states are considered distinguishable if they are orthogonal. In this case, a measurement exists that is guaranteed to determine which of the two states a system is in.
- Efficient computation.** A computation is efficient if it requires, at most, polynomially many resources as a function of input size. For example, if the computation returns the value $f(x)$ on input x , where x is a bit string, then it is efficient if there exists a power k such that the number of computational steps used to obtain $f(x)$ is bounded by $|x|^k$, where $|x|$ is the length (number of bits) of x .
- Entanglement.** A nonclassical correlation between two quantum systems most strongly exhibited by the maximally entangled states, such as the Bell states for two qubits, and considered to be absent in mixtures of product states (which are called separable states). Often, states that are not separable are considered to be entangled. However, nearly separable states do not exhibit all the features of maximally entangled states. As a result, studies of different types of entanglement are an important component of quantum information theory.
- Gate.** An operation applied to information for the purpose of information processing.

Global phase. Two quantum states are indistinguishable if they differ only by a global phase. That is, $|\psi\rangle$ and $e^{i\phi}|\psi\rangle$ are in essence the same state. The global phase difference is the factor $e^{i\phi}$. The equivalence of the two states is apparent from the fact that their density matrices are the same.

Hilbert space. An n -dimensional Hilbert space consists of all complex n -dimensional vectors. A defining operation in a Hilbert space is the inner product. If the vectors are thought of as column vectors, then the inner product $\langle x, y \rangle$ of x and y is obtained by forming the conjugate transpose x^\dagger of x and calculating $\langle x, y \rangle = x^\dagger y$. The inner product induces the usual squared norm $|x|^2 = \langle x, x \rangle$.

Information. Something that can be recorded, communicated, and computed with. Information is fungible; that is, its meaning can be identified regardless of the particulars of the physical realization. Thus, information in one realization (such as ink on a sheet of paper) can be easily transferred to another (for example, spoken words). Types of information include deterministic, probabilistic, and quantum information. Each type is characterized by information units, which are abstract systems whose states represent the simplest information of each type. The information units define the “natural” representation of the information. For deterministic information, the information unit is the bit, whose states are symbolized by 0 and 1. Information units can be put together to form larger systems and can be processed with basic operations acting on few of them at a time.

Inner product. The defining operation of a Hilbert space. In a finite dimensional Hilbert space with a chosen orthonormal basis $\{e_i : 1 \leq i \leq n\}$, the inner product of two vectors $x = \sum_i x_i e_i$ and $y = \sum_i y_i e_i$ is given by $\sum_i \bar{x}_i y_i$. In the standard column representation of the two vectors, this is the number obtained by computing the product of the conjugate transpose of x with y . For real vectors, that product agrees with the usual “dot” product. The inner product of x and y is often written in the form $\langle x, y \rangle$. Pure quantum states are unit vectors in a Hilbert space. If $|\phi\rangle$ and $|\psi\rangle$ are two quantum states expressed in the ket-bra notation, their inner product is given by $(|\phi\rangle)^\dagger |\psi\rangle = \langle \phi | \psi \rangle$.

Ket. A state expression of the form $|\psi\rangle$ representing a quantum state. Usually, $|\psi\rangle$ is thought of as a superposition of members of a logical state basis $|i\rangle$. One way to think about the notation is to consider the two symbols $|$ and \rangle as delimiters denoting a quantum system and ψ as a symbol representing a state in a standard Hilbert space. The combination $|\psi\rangle$ is the state of the quantum system associated with ψ in the standard Hilbert space via a fixed isomorphism. In other words, one can think of $\psi \leftrightarrow |\psi\rangle$ as an identification of the quantum system’s state space with the standard Hilbert space.

Linear extension of an operator. The unique linear operator that implements a map defined on a basis. Typically, we define an operator U on a quantum system only on the logical states $U : |i\rangle \rightarrow |\psi_i\rangle$. The linear extension is defined by $U(\sum_i \alpha_i |i\rangle) = \sum_i \alpha_i |\psi_i\rangle$.

Logical states. For quantum systems used in information processing, the logical states are a fixed orthonormal basis of pure states. By convention, the logical basis for qubits consists of $|0\rangle$ and $|1\rangle$. For larger dimensional quantum systems, the logical basis is often indexed by integers, $|0\rangle$, $|1\rangle$, $|2\rangle$, and so on. The logical basis is often called the computational basis, or sometimes, the classical basis.

Measurement. The process used to extract classical information from a quantum system. A general projective measurement is defined by a set of projectors P_i , satisfying $\sum_i P_i = \mathbb{1}$ and $P_i P_j = \delta_{ij} P_i$. Given the quantum state $|\psi\rangle$, the outcome of a measurement with the set $\{P_i\}_i$ is one of the classical indices i associated with a projector P_i . The index i is the measurement outcome. The probability of outcome i is $p_i = |P_i |\psi\rangle|^2$, and given outcome i , the quantum state “collapses” to $P_i |\psi\rangle / \sqrt{p_i}$.

Mixture. A probabilistic combination of the pure states of a quantum system. Mixtures can be represented without redundancy with density operators. Thus, a mixture is of the form $\sum_i \lambda_i |\psi_i\rangle\langle\psi_i|$, with $\lambda_i \geq 0$ and $\sum_i \lambda_i = 1$ being the probabilities of the states $|\psi_i\rangle$. This expression for mixtures defines the set of density operators, which can also be characterized as the set of operators ρ satisfying $\text{tr}(\rho) = 1$ and for all $|\psi\rangle$, $\langle\psi|\rho|\psi\rangle \geq 0$ (“positive semidefinite operator”).

Network. In the context of information processing, a network is a sequence of gates applied to specified information units. Networks can be visualized as displaying horizontal lines that denote the timeline of an information unit. The gates are represented by graphical elements that intercept the lines at specific points. A realization of the network requires applying the gates to the information units in the specified order (left to right).

Operator. A function that transforms the states of a system. Operators may be restricted depending on the system’s properties. For example, in talking about operators acting on quantum systems, one always assumes that they are linear.

Oracle. An information processing operation that can be applied. A use of the oracle is called a query. In the oracle model of computation, a standard model is extended to include the ability to query an oracle. Each oracle query is assumed to take one time unit. Queries can reduce the resources required for solving problems. Usually, the oracle implements a function or solves a problem not efficiently implementable by the model without the oracle. Oracle models are used to compare the power of two models of computation when the oracle can be defined for both models. In 1994, for example, Dan Simon showed that quantum computers with a specific oracle O could efficiently solve a problem that had no efficient solution on classical computers with access to the classical version of O . At the time, this result was considered the strongest evidence for an exponential gap in power between classical and quantum computers.

Overlap. The inner product between two quantum states.

Pauli operators. The Hermitian matrices σ_x , σ_y , and σ_z acting on qubits, which are two-level quantum systems. They are defined in Equation (12). It is often convenient to consider the identity operator to be included in the set of Pauli operators.

Polynomial resources. To say that an algorithm computing the function $f(x)$, where x is a bit string, uses polynomial resources (in other words, is efficient) means that the number of steps required to compute $f(x)$ is bounded by $|x|^k$ for some fixed k . Here, $|x|$ denotes the length of the bit string x .

Probabilistic bit. The basic unit of probabilistic information whose state space consists of all probability distributions over the two states of a bit. The states can be thought of as describing the outcome of a biased coin flip before the coin is flipped.

Probabilistic information. The type of information obtained by extending the state spaces of deterministic information to allow arbitrary probability distributions over the deterministic states. This is the main type of classical information with which quantum information is compared.

Probability amplitude. The squared norm of an amplitude with respect to a chosen orthonormal basis $\{|i\rangle\}$. Thus, the probability amplitude is the probability with which the state $|i\rangle$ is measured in a complete measurement that uses this basis.

Product state. For two quantum systems A and B, product states are of the form $|\psi\rangle_A |\phi\rangle_B$. Most states are not of this form.

Program. An algorithm expressed in a language that can be understood by a particular type of computer.

Projection operator. A linear operator P on a Hilbert space that satisfies $P^2 = P^\dagger P = P$. The projection onto a subspace V with orthogonal complement W is defined as follows: If $x \in V$ and $y \in W$, then $P(x + y) = x$.

Pseudocode. A semiformal computer language intended to be executed by a standard random-access machine, which is a machine model with a central processing unit and access to a numerically indexed unbounded memory. This machine model is representative of the typical one-processor computer. Pseudocode is similar to programming languages such as BASIC, Pascal, or C but does not have specialized instructions for human interfaces, file management, or other “external” devices. Its main use is to describe algorithms and enable machine-independent analysis of the algorithms’ resource usage.

Pure state. A state of a quantum system that corresponds to a unit vector in the Hilbert space used to represent the system’s state space. In the ket notation, pure states are written in the form $|\psi\rangle = \sum_i \alpha_i |i\rangle$, where the $|i\rangle$ form a logical basis and $\sum_i |\alpha_i|^2 = 1$.

Quantum information. The type of information obtained when the state space of deterministic information is extended by normalized superpositions of deterministic states. Formally, each deterministic state is identified with one of an orthonormal basis vector in a Hilbert space, and normalized superpositions are unit-length vectors expressible as complex linear sums of the chosen basis vectors. It is convenient to extend this state space further by permitting probability distributions over the quantum states (see the entry for “mixtures”). This extension is still called quantum information.

Qubit. The basic unit of quantum information. It is the quantum extension of the deterministic bit, which implies that its state space consists of the unit-length vectors in a two-dimensional Hilbert space.

Readout. A method for obtaining human-readable information from the state of a computer. For quantum computers, readout refers to a measurement process used to obtain classical information about a quantum system.

Reversible gate. A gate whose action can be undone by a sequence of gates.

Separable state. A mixture of product states.

States. The set of states for a system characterizes the system’s behavior and possible configurations.

Subspace. For a Hilbert space, a subspace is a linearly closed subset of the vector space.

The term can be used more generally for a system Q of any information type:

A subspace of Q or, more specifically, of the state space of Q is a subset of the state space that preserves the properties of the information type represented by Q .

Superposition principle. One of the defining postulates of quantum mechanics according to which if states $|1\rangle, |2\rangle, \dots$ are distinguishable, then $\sum_i \alpha_i |i\rangle$ with $\sum_i |\alpha_i|^2 = 1$ is a valid quantum state. Such a linear combination is called a normalized superposition of the states $|i\rangle$.

System. An entity that can be in any of a specified number of states. An example is a desktop computer whose states are determined by the contents of its various memories and disks. Another example is a qubit, which can be thought of as a particle whose state space is identified with complex, two-dimensional, length-one vectors. Here, a system is always associated with a type of information that determines the properties of the state space. For example, for quantum information, the state space is a Hilbert space. For deterministic information, it is a finite set called an alphabet.

Unitary operator. A linear operator U on a Hilbert space that preserves the inner product. That is, $\langle Ux, Uy \rangle = \langle x, y \rangle$. If U is given in matrix form, then this expression is equivalent to $U^\dagger U = \mathbb{1}$.

Universal set of gates. A set of gates that satisfies the requirement that every allowed operation on information units can be implemented by a network of these gates. For quantum information, it means a set of gates that can be used to implement every unitary operator. More generally, a set of gates is considered universal if, for every operator U , there are implementable operators V arbitrarily close to U .

From Factoring to Phase Estimation

A discussion of Shor's algorithm

Emanuel Knill, Raymond Laflamme, Howard N. Barnum, Diego A. Dalvit, Jacek J. Dziarmaga, James E. Gubernatis, Leonid Gurvits, Gerardo Ortiz, Lorenza Viola, and Wojciech H. Zurek

The publication of Shor's quantum algorithm for efficiently factoring numbers (1994 and 1997) was the key event that stimulated many theoretical and experimental investigations of quantum computation. One of the reasons why this algorithm is so important is that the security of widely used public-key cryptographic protocols relies on the conjectured difficulty of factoring large numbers. An elementary overview of these protocols and the quantum algorithm for breaking them is provided in Artur Ekert (1998).¹ Here, we outline the relationship between factoring and the powerful technique of phase estimation. This relationship helps in understanding many of the existing quantum algorithms and was first explained in Richard Cleve et al. (1998). This explanation was motivated by Alexei Kitaev's version (1995) of the factoring algorithm.

The factoring problem requires writing a whole number N as a product of primes. (Primes are whole numbers greater than 1 that are divisible without remainder only by 1 and themselves.) Shor's algorithm solves this problem by reducing it to instances of the order-finding problem, which will be defined below. The reduction is based on basic number theory and involves efficient classical computation. At the core of Shor's algorithm is a quantum algorithm that solves the order-finding problem efficiently. In this case, an algorithm is considered efficient if it uses resources bounded by a polynomial in the number of digits of N . For more information on the requisite number theory, see any textbook on number theory (Bolker 1970, Hardy and Wright 1979).

We begin by showing that factoring reduces to order finding. The first observation is that, to factor a whole number, it is sufficient to solve the factor-finding problem, whose statement is, "Given a whole number N , find a proper factor of N if one exists. A factor of N is a whole number f that satisfies $N = fg$ for some whole number g . The factor f is proper if $f \neq 1$ and $f \neq N$. For example, if $N = 15$, then 3 and 5 are its proper factors. For some numbers, it is easy to find proper factors. For example, you can tell that N is even from the least significant digit (in decimal or binary), in which case, 2 is a proper factor (unless $N = 2$, a prime). But many numbers are not so easy. As an example, you can try to find a factor of $N = 149,573$ by hand.² You can complete the factorization of a whole number by recursively applying an algorithm for the factor-finding problem to all the proper factors found.

Before we continue the reduction of factoring to order finding, we will briefly explain modular arithmetic, which both simplifies the discussion and is necessary to avoid computing with numbers that have exponential numbers of digits. We say that a and b are equal modulo N , written as $a = b \bmod N$, if $a - b$ is divisible by N (without remainder). For example, $3 = 18 \bmod 15 = 33 \bmod 15$. Equality modulo N is well behaved with respect to addition and multiplication. That is, if $a = b \bmod N$ and $c = d \bmod N$, then $a + c = b + d \bmod N$, and $ac = bd \bmod N$. For factoring N , we will be look-

¹All the citations in this article have been referenced on [pages 31 to 33](#) of the main article, "Quantum Information Processing."

² $149573 = 373 \times 401$

ing for whole numbers a that are divisible by a proper factor of N . If a has this property, then so does any b with $b = a \bmod N$. We therefore perform all arithmetic modulo N .

One way to think of all this is that we use only whole numbers a that satisfy $0 \leq a \leq N - 1$. We can implement each arithmetic operation modulo N by applying the operation in the usual way and then computing the remainder after division by N . For example, to obtain $ab \bmod N$, we first compute ab . The unique c such that $0 \leq c \leq N - 1$ and $c = ab \bmod N$ is the remainder after division of ab by N . Thus, c is the result of multiplying a by b modulo N . Consistent with this procedure, we can think of the expression $a \bmod N$ as referring to the remainder of a after division by N .

The second observation in the reduction of factoring to order finding is that it is sufficient to find a whole number r with the property that $r^2 - 1$ is a multiple of N , but $r - 1$ and $r + 1$ are not. Using the language of modular arithmetic, the property is expressed as $r^2 = 1 \bmod N$, but $r \neq 1 \bmod N$ and $r \neq -1 \bmod N$. Because $1 \bmod N$ and $-1 \bmod N$ are the obvious square roots of $1 \bmod N$, we say that r is a nontrivial square root of unity (modulo N). For such an r , one can write $r^2 - 1 = (r - 1)(r + 1) = mN$ for some whole number m . This implies that every prime factor p of N divides either $(r - 1)$ or $(r + 1)$ so that either $(r - 1)$ or $(r + 1)$ is or shares a factor with N . Suppose that $r - 1$ is or shares such a factor. Because $r - 1$ is not a multiple of N , the greatest common divisor of $r - 1$ and N is a factor of N . Since an efficient classical algorithm (the Euclidean algorithm) exists for finding the greatest common divisor, we can easily find the desired proper factor.

The examples of $N = 15$ and $N = 21$ serve to illustrate the key features of the algorithm. For $N = 15$, possible choices for r are $r = 4$ ($4^2 - 1 = 1 * 15$), and $r = 11$ ($11^2 - 1 = 120 = 8 * 15$). For the first choice, the proper factors emerge immediately: $4 - 1 = 3$, and $4 + 1 = 5$. For the second, it is necessary to determine the greatest common divisors (or gcd). Let $\gcd(x, y)$ stand for the greatest common divisor of x and y . The proper factors are $\gcd(11 - 1, 15) = \gcd(10, 15) = 5$, and $\gcd(11 + 1, 15) = \gcd(12, 15) = 3$. For $N = 21$, one can take $r = 8$ as $8^2 - 1 = 63 = 3 * 21$. In this case, $8 - 1 = 7$ is a proper factor, and $\gcd(8 + 1, 21) = 3$ is another.

For N even or a power of a prime, it is not always possible to find a nontrivial square root of unity. Because both cases can be handled efficiently by known classical algorithms, we can exclude them. In every other case, such numbers r exist. One way to find such an r is to start from any whole number q , with $1 < q < N$. If $\gcd(q, N) = 1$, then according to a basic result in number theory, there is a smallest whole number $k > 1$ such that $q^k - 1 = 0 \bmod N$. The number k is called the order of q modulo N . If k is even, say, $k = 2l$, then $(q^l)^2 = 1 \bmod N$, so q^l is a (possibly trivial) square root of unity. For the example of $N = 15$, we can try $q = 2$. The order of 2 modulo 15 is 4, which gives $r = 2^2 = 4$, the first of the two choices in the previous paragraph. For $N = 21$, again with $q = 2$, the order is 6: $2^6 - 1 = 63 = 3 * 21$. Thus, $r = 2^3 = 8$. We can also try $q = 11$, in which case, with foresight, it turns out that $11^6 - 1$ is divisible by 21. A possible problem appears, namely, the powers q^k , which we want to compute, are extremely large. But modular arithmetic can help us avoid this problem. For example, to find the order of 11 modulo 21 by direct search, we can perform the following computation: In general, such a direct search for the order of q modulo N is very inefficient, but as we will see,

$$\begin{aligned}
 11^2 &= 121 = 5 * 21 + 16 = 16 \bmod 21 \\
 11^3 &= 11 * 11^2 = 11 * 16 \bmod 21 = 11 * (-5) \bmod 21 \\
 &= -55 \bmod 21 = -3 * 21 + 8 \bmod 21 = 8 \bmod 21 \\
 11^4 &= 11 * 11^3 = 11 * 8 \bmod 21 = 4 * 21 + 4 \bmod 21 = 4 \bmod 21 \\
 11^5 &= 11 * 11^4 = 11 * 4 \bmod 21 = 2 \bmod 21 \\
 11^6 &= 11 * 11^5 = 11 * 2 \bmod 21 = 1 \bmod 21
 \end{aligned} \tag{1}$$

there is an efficient quantum algorithm that can determine the order.

A factor-finding algorithm based on the above observations is the following:

FACTORFIND(N)

Input: A positive, nonprime whole number N

Output: A proper factor f of N , that is, f is a whole number such that $1 < f < N$ and $N = fg$ for some whole number g .

1. If N is even, return $f = 2$.
2. If $N = p^k$ for p prime, return p .
3. Randomly pick $1 < q < N - 1$.
 - a. If $f = \gcd(q, N) > 1$, return f .
4. Determine the order k of q modulo N using the quantum order-finding algorithm.
 - a. If k is not even, repeat at step 3.
5. Write $k = 2l$ and determine $r = q^l \bmod N$ with $1 < r < N$.
 - a. If $1 < f = \gcd(r - 1, N) < N$, return f .
 - b. If $1 < f = \gcd(r + 1, N) < N$, return f .
 - c. If we failed to find a proper factor, repeat at step 3.

The efficiency of this algorithm depends on the probability that a randomly chosen q at step 3 results in finding a factor. An analysis of the group of numbers q that satisfy $\gcd(q, N) = 1$ shows that this probability is sufficiently large.

The main problem left to be solved is finding the order of $q \bmod N$. A direct search for the order of $q \bmod N$ involves computing the sequence

$$1 \rightarrow q \rightarrow q^2 \bmod N \rightarrow \dots \rightarrow q^{k-1} \bmod N \rightarrow 1 = q^k \bmod N . \quad (2)$$

This sequence can be conveniently visualized as a cycle whose length is the order $q \bmod N$ (refer to Figure 1).

To introduce the quantum algorithm, we first associate the logical quantum states $|0\rangle, |1\rangle, \dots, |N-1\rangle$ with the numbers $0, 1, \dots, N-1$. The map f that takes each number on the cycle to the next number along the cycle is given by $f(x) = qx \bmod N$. For q satisfying $\gcd(q, N) = 1$, the map f permutes not only the numbers on the cycle but all the numbers modulo N . As a result, the linear operator \hat{f} defined by $\hat{f}|x\rangle = |f(x)\rangle = |qx \bmod N\rangle$ is unitary. The quantum algorithm deduces the length of the cycle for q by making measurements to determine the properties of the action of \hat{f} on superpositions of the states $|q^s \bmod N\rangle$. To illustrate the basic ideas, we work out the example of $N = 15$ and $q = 8$. The action of \hat{f} on the states $|1\rangle, |8\rangle, |4\rangle$, and $|2\rangle$ in the cycle of $8 \bmod 15$ is

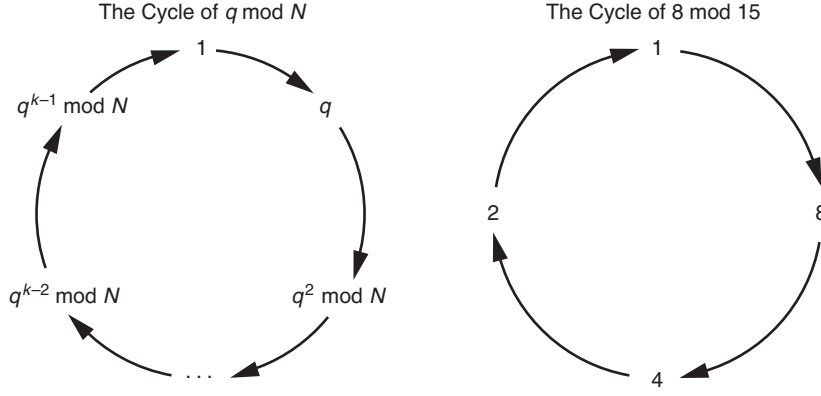


Figure 1. Multiplicative Cycles of $q \bmod N$
Each number on a cycle is obtained from the previous one by multiplication by $q \bmod N$.

completely determined by the eigenstates and eigenvalues of \hat{f} . For cyclicly acting permutations, a basis of eigenstates is given by the Fourier basis for the space spanned by the states in a cycle. For the cycle of interest, the Fourier basis consists of the states

$$\begin{aligned} |\psi_0\rangle &= \frac{1}{2}(|1\rangle + |8\rangle + |4\rangle + |2\rangle), \\ |\psi_1\rangle &= \frac{1}{2}(|1\rangle + i|8\rangle - |4\rangle - i|2\rangle), \\ |\psi_2\rangle &= \frac{1}{2}(|1\rangle - |8\rangle + |4\rangle - |2\rangle), \text{ and} \\ |\psi_3\rangle &= \frac{1}{2}(|1\rangle - i|8\rangle - |4\rangle + i|2\rangle). \end{aligned} \quad (3)$$

The phases of the l^{th} state of the cycle occurring in the sum for $|\psi_m\rangle$ can be written as i^{lm} . It follows that $\hat{f}|\psi_m\rangle = i^m|\psi_m\rangle$, that is, the eigenvalue of \hat{f} for $|\psi_m\rangle$ is i^m . Note that, in complex numbers, the powers of i are all the fourth roots of unity. In general, the Fourier basis for the cycle $\dots \rightarrow |q^l \bmod N\rangle \rightarrow \dots$ consists of the states $|\psi_m\rangle = \sum_l \omega^{lm} |q^l \bmod N\rangle$, where $\omega = e^{i2\pi/k}$ is a primitive k^{th} root of unity. (The complex number x is a primitive k^{th} root of unity if k is the smallest whole number $k > 0$ such that $x^k = 1$. For example, both -1 and i are fourth roots of unity, but only i is primitive.)

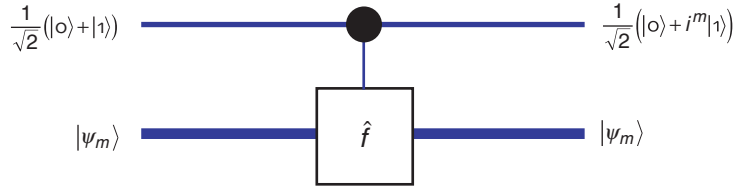
It is, of course, possible to express the logical state $|1\rangle$ using the Fourier basis

$$|1\rangle = \frac{1}{2}(|\psi_0\rangle + |\psi_1\rangle + |\psi_2\rangle + |\psi_3\rangle). \quad (4)$$

The key step of the quantum algorithm for order finding consists of a measurement to estimate a random eigenvalue of \hat{f} , whose associated eigenstate occurs in the expression for $|1\rangle$ in terms of the Fourier basis. If the eigenvalue found is a k^{th} root of unity, we infer that the cycle length is divisible by k and check (using a classical algorithm) whether this is the order of q . In the example, the random eigenvalues are 1 (the only primitive first root of unity), i and $-i$ (primitive fourth roots of unity), and -1 (the primitive second root of unity). The order is found if the random eigenvalue is a fourth root of unity, which happens with probability $1/2$ in this case.

Figure 2. Phase Estimation with One Qubit

The input is a product state on one ancilla qubit and on a second quantum system, as shown. The state $|\psi_m\rangle$ on the second system is an eigenstate of \hat{f} . For the example provided in Equation (3), the eigenvalue is i^m . A controlled- \hat{f} operation is applied to the input, that is, \hat{f} is applied to the second system conditional on $|1\rangle$ for the ancilla qubit. In the bra-ket notation, the total operation can be written as $|0\rangle\langle 0| + |1\rangle\langle 1|\hat{f}$ (system labels have been omitted). Because \hat{f} changes only the phase of its input, the second system is unchanged, but the phase modifies the ancilla qubit's superposition as shown.



The quantum algorithm for obtaining an eigenvalue is called the phase estimation algorithm, and it exploits a more general version of the phase kickback we encountered in the solution of the parity problem. The phase kickback transfers the eigenvalue of an eigenstate of \hat{f} to a Fourier basis on a number of additional qubits called helper or ancilla qubits. Which Fourier state results is then determined by a subroutine called the measured quantum Fourier transform. We introduce these elements in the next paragraphs. Their combination for solving the general order-finding problem is illustrated on [page 45](#).

Figure 2 shows how to kick back the eigenvalue of an eigenstate of \hat{f} using a network implementing the controlled- \hat{f} operation. The network in Figure 2 can be used with input $|1\rangle$ on the second system. From Equation (4) and the superposition principle, it follows that the output correlates the different phase kickback states with the four eigenvectors $|\psi_m\rangle$. That is, the network implements the following transformation:

$$\frac{1}{2\sqrt{2}} (|0\rangle + |1\rangle) \begin{pmatrix} |\psi_0\rangle \\ + |\psi_1\rangle \\ + |\psi_2\rangle \\ + |\psi_3\rangle \end{pmatrix} \rightarrow \frac{1}{2\sqrt{2}} \begin{pmatrix} (|0\rangle + i^0|1\rangle) |\psi_0\rangle \\ + (|0\rangle + i^1|1\rangle) |\psi_1\rangle \\ + (|0\rangle + i^2|1\rangle) |\psi_2\rangle \\ + (|0\rangle + i^3|1\rangle) |\psi_3\rangle \end{pmatrix}. \quad (5)$$

The hope is that a measurement of the first qubit can distinguish between the four possible phases that can be kicked back. However, because the four states are not mutually orthogonal, they are not unambiguously distinguishable by a measurement. To solve this problem, we use a second qubit and a controlled- \hat{f}^2 as shown in Figure 3.

The four possible states $|u_m\rangle$ that appear on the ancilla qubits in the network of Figure 3 are the Fourier basis for the cycle $0 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow 0$ and are therefore orthonormal. If we apply the network of Figure 3 with $|1\rangle$ instead of $|\psi_m\rangle$ at the lower input, the output correlates the four $|\psi_m\rangle$ in the superposition with the $|u_m\rangle$, which makes the information about the eigenvalues of \hat{f} available in the Fourier basis of the two ancilla qubits. This approach has the advantage that the states are known, whereas in the Fourier basis for the cycle of $q \bmod N$, the states depend on the numbers in the cycle, which are not known in advance (except in very simple cases, such as the example we are working with).

To learn one of the eigenvalues of \hat{f} , the last step is to make a measurement in the Fourier basis. For one qubit representing the binary numbers 0 and 1, the Fourier basis is $1/\sqrt{2}(|0\rangle + |1\rangle)$ and $1/\sqrt{2}(|0\rangle - |1\rangle)$, which is constructed as discussed after

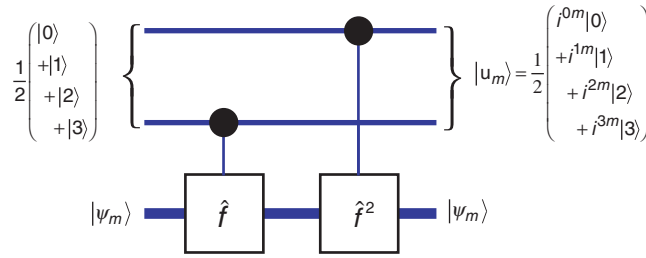


Figure 3. Phase Estimation with Two Qubits

Using two qubits ensures distinguishability of the eigenvalues of \hat{f} for the states $|\psi_m\rangle$. The states of the input qubits are used to represent the numbers from 0 to 3 in binary. The most significant bit (the two's digit in binary representation) is carried by the top qubit. That is, we make the following identification: $|0\rangle = |00\rangle$, $|1\rangle = |01\rangle$, $|2\rangle = |10\rangle$, and $|3\rangle = |11\rangle$. It follows that the network has the effect of applying \hat{f}^m conditional on the input qubits' logical state being $|m\rangle$.

Equation (3) but using the square root of unity $\omega = -1$ instead of the fourth root i . To make a measurement that determines which of the two basis vectors is present, it suffices to apply the Hadamard transform \mathbf{H} and make a standard measurement, just as we did twice in the network of Figure 2 in the article “Quantum Information Processing” on page 23. A more complicated network works with two qubits representing the binary numbers from 0 to 3. Such a network is shown in Figure 4.

To see how the network extracts the bits in the index of $|u_a\rangle$, we can follow the states as the network is executed. The input state at checkpoint 1 in Figure 4 is given by

$$|\phi_1\rangle = |u_a\rangle = \frac{1}{2} \begin{pmatrix} i^{0*a}|0\rangle \\ +i^{1*a}|1\rangle \\ +i^{2*a}|2\rangle \\ +i^{3*a}|3\rangle \end{pmatrix} = \frac{1}{2} \begin{pmatrix} i^{(0*2^1+0*2^0)}(a_1*2^1+a_0*2^0)|00\rangle \\ +i^{(0*2^1+1*2^0)}(a_1*2^1+a_0*2^0)|01\rangle \\ +i^{(1*2^1+0*2^0)}(a_1*2^1+a_0*2^0)|10\rangle \\ +i^{(1*2^1+1*2^0)}(a_1*2^1+a_0*2^0)|11\rangle \end{pmatrix}. \quad (6)$$

In the last sum, the relevant numbers have been fully expanded in terms of their binary digits to give a flavor of the general principles underlying the measured Fourier transform. The next step of the network applies a Hadamard gate to the qubit carrying the most significant digit. To understand how it succeeds in extracting a_0 , the least signifi-

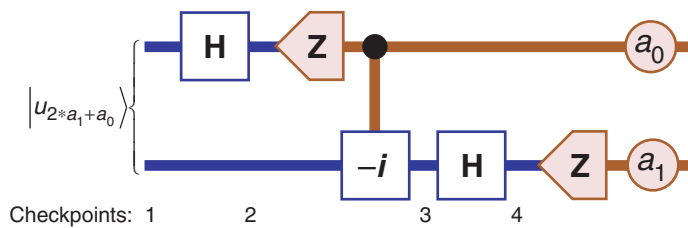


Figure 4. Measured Quantum Fourier Transform on Two Qubits
The two qubits represent the numbers 0, 1, 2, and 3. If the input is one of the Fourier states $|u_a\rangle$, where the binary digits of a are determined by $a = 2 * a_1 + a_0$, then the measurement outcomes are a_0 and a_1 , as shown. The numbers under the network are checkpoints used for analysis. [For details on the measured Fourier transform, see Griffiths and Niu (1996).]

cant bit of a , let b with binary digits b_0 and b_1 represent one of the logical states of the two qubits. As before, the most significant bit b_1 is represented by the top/first qubit that the first Hadamard gate is applied to. The phase of $|b\rangle$ in Equation (6) is given by $i^{(b_1*2^1+b_0*2^0)(a_1*2^1+a_0*2^0)}$. Next, we determine how the phase depends on b_1 :

$$\begin{aligned}
 i^{(b_1*2^1+b_0*2^0)(a_1*2^1+a_0*2^0)} &= i^{b_1*2^1*(a_1*2^1+a_0*2^0)} i^{b_0*2^0*(a_1*2^1+a_0*2^0)} \\
 &= i^{b_1*a_1*2^2} i^{b_1*a_0*2^1} i^{b_0*2^0*(a_1*2^1+a_0*2^0)} \\
 &= (i^4)^{b_1*a_1} (i^2)^{b_1*a_0} i^{b_0*2^0*(a_1*2^1+a_0*2^0)} \\
 &= (-1)^{b_1*a_0} i^{b_0*2^0*(a_1*2^1+a_0*2^0)} .
 \end{aligned} \tag{7}$$

It follows that, if $a_0 = 0$, the phase does not depend on b_1 , and if $a_0 = 1$, it changes sign with b_1 . This sign change can be detected by performing the Hadamard transform and measuring, as can be seen explicitly by computing the state after the Hadamard transform at checkpoint 2:

$$\begin{aligned}
 |\phi_2\rangle &= \frac{1}{\sqrt{2}} \left(i^{0*2^0*(a_1*2^1+a_0*2^0)} |a_0\rangle |0\rangle + i^{1*2^0*(a_1*2^1+a_0*2^0)} |a_0\rangle |1\rangle \right) \\
 &= |a_0\rangle \frac{1}{\sqrt{2}} \left(i^{0*2^0*(a_1*2^1+a_0*2^0)} |0\rangle + i^{1*2^0*(a_1*2^1+a_0*2^0)} |1\rangle \right) .
 \end{aligned} \tag{8}$$

The phases still show a dependence on a_0 via the terms $i^{b_0*2^0*a_0*2^0} = i^{b_0a_0}$. The purpose of the phase-shift gate conditioned on the measurement outcome is to remove that dependence. The result is the following state on the remaining qubit at checkpoint 3:

$$\begin{aligned}
 |\phi_3\rangle &= \frac{1}{\sqrt{2}} \left(i^{0*2^0*a_1*2^1} |0\rangle + i^{1*2^0*a_1*2^1} |1\rangle \right) \\
 &= \frac{1}{\sqrt{2}} \left((-1)^{0*a_1} |0\rangle + (-1)^{1*a_1} |1\rangle \right) \\
 &= \frac{1}{\sqrt{2}} \left(|0\rangle + (-1)^{a_1} |1\rangle \right) .
 \end{aligned} \tag{9}$$

The final Hadamard transform followed by a measurement therefore results in the bit a_1 , as desired.

The elements that we used to determine the order of 8 modulo 15 can be combined and generalized to determine the order of any q modulo N with $\gcd(q, N) = 1$. The general network is shown in Figure 5. Two features of the generalization are not apparent from the example. First, in order for the quantum network to be efficient, an efficient implementation of the controlled \hat{f}^{2^l} operation is required. To obtain such an implementation, first note that to calculate $\hat{f}^{2^l}(x) = q^{2^l}x \bmod N$, it suffices to square q repeatedly

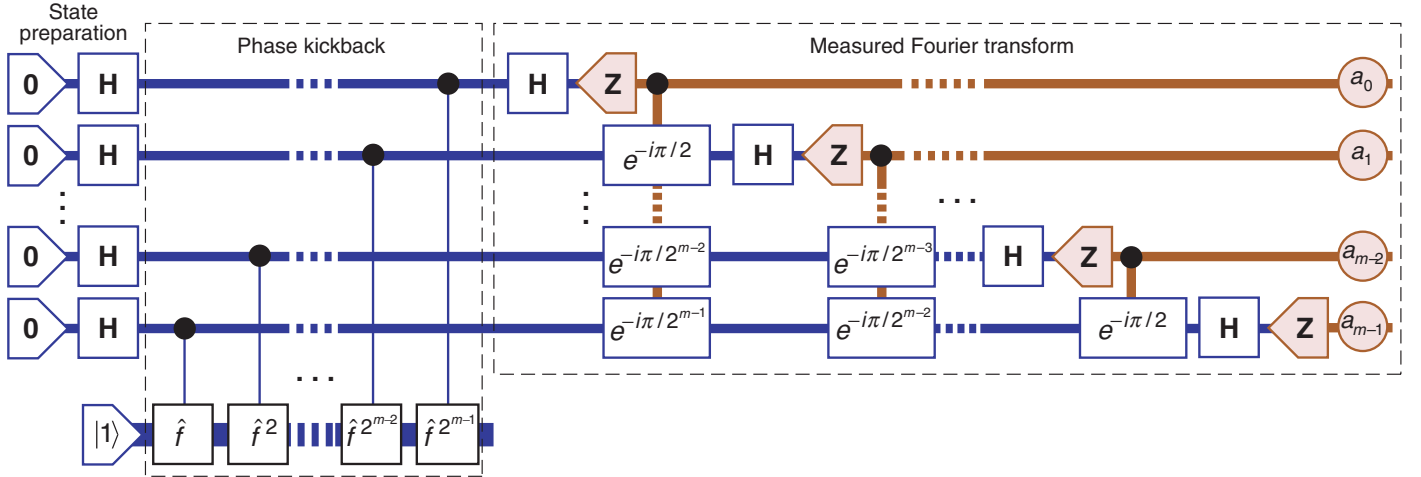


Figure 5. Network for Quantum Order Finding and Phase Estimation

The number m of qubits used for the phase kickback has to be chosen such that $m > 2 * \log_2(k_u)$, where k_u is a known upper bound on the order k of $q \bmod N$. Because $N > k$, one can set $m = 2 \lceil \log_2(N) \rceil$, where $\lceil x \rceil$ is the least whole number $s \geq x$. There is an eigenvalue $\lambda_l = e^{i2\pi l/k}$ of one of the Fourier eigenvectors associated with the cycle of $q \bmod N$ such that the number a , whose binary digits are the measurement outcomes, satisfies $e^{i\pi a/2^{m-1}} \approx e^{i2\pi l/k}$. More precisely, with probability above .405, there exists l such that $|a/2^m - l/k| \leq 1/2^{m+1}$ (Cleve et al. 1998). Because any two distinct rational numbers with denominator at most k_u differ by at least $1/k_u^2 > 2/2^{m+1}$, the theory of rational

approximations guarantees that we can uniquely determine the number l/k . There is an efficient classical algorithm based on continued fractions that computes r and s with $r/s = l/k$ and $s = k/\gcd(l, k)$. The probability that $\gcd(l, k) = 1$ is at least $1/(\log_2(k) + 1)$, in which case we learn that $s = k$ and this is the order of $q \bmod N$. Note that the complexity of the network depends on the complexity of implementing the controlled \hat{f}^{2^l} operations. Because these operations can be implemented efficiently, the network and hence the determination of the order of $q \bmod N$ are efficient in the sense that, on average, polynomial resources in $\log_2(N)$ suffice.

modulo N using $(q^{2^m})^2 \bmod N = q^{2^{m+1}} \bmod N$ until we obtain $q^{2^l} \bmod N$. The result is then multiplied by $x \bmod N$. This computation is efficient. For any given q , the computation can be converted to an efficient network consisting of Toffoli gates and controlled-not gates acting on the binary representation of x . The conversion can be accomplished with standard techniques from the theory of reversible classical computation. The result is an efficient network for \hat{f}^{2^l} . Basic network theory can then be used to implement the controlled version of this operation (Barenco et al. 1995).

To understand the second feature, note that we were lucky to anticipate that the order of 8 modulo 15 was a power of 2, which nicely matched the measured Fourier transform we constructed on two qubits. The measured Fourier transform on m ancilla qubits can detect exactly only eigenvalues that are powers of the 2^m th root of unity $e^{i\pi/2^{m-1}}$. The phase kicked back by the controlled operations corresponds to a k^{th} root of unity. Given a Fourier state on the cycle of $q \bmod N$, the resulting state on the ancilla qubits has phases that go as powers of a k^{th} root of unity. Fortunately, the ancilla's Fourier basis is such that the measured Fourier transform picks up primarily those basis states whose generating phase is close to the kickback phase. Thus, we are likely to detect a nearby $\omega = e^{i\pi/2^{m-1}}$. It is still necessary to infer (a divisor of) k from knowledge of such an ω . Because we know that the order k is bounded by N , the number of possible phases kicked back that are near the measured ω is limited. To ensure that there is only one possible such phase, it is necessary to choose m such that $2^m > N^2$. (See also Figure 5.) ■



cryptography
quantum computers
cryptography
quantum computers

QUESTIONS, Quantum Computers, and Cryptography

A mathematical metaphor for the power of quantum algorithms

Mark Ettinger

How can quantum computers do the amazing things that they are able to do, such as factoring large numbers and finding discrete logarithms? What makes them so different from classical computers? These questions are often asked, and they have proved to be surprisingly difficult to answer—at least to the satisfaction of everyone! In this short article, I'll try to address these questions by comparing the operation of a quantum computer with playing the game of 20 questions. But first, let's consider an unusual perspective on computers in general.

What Is a Computer?

Well, a computer is really just some physical machine that you prepare in a certain way, manipulate in certain ways, and then watch to observe the results it displays. That is how physicists might describe the entire physical process that mathematicians call a computation. This view seems a bit strange at first because we have become accustomed to the more abstract view of the computer scientist, who sees a computation as a certain type of process that acts on an input in order to produce an output. But our physical description is not really so different. It just emphasizes the physical nature of the computation, something that falls by the wayside in the abstracted view. The initial preparation is what a computer scientist calls an input, the actual computation is the physical manipulation, and the observation at the end results in getting the output. So, whereas a computation can be viewed abstractly as a process, its physical nature can also be emphasized. This view will help us make the transition to understanding what a quantum machine is doing in a special way. Unlike classical computers, which are physical devices manipulated according to the laws of classical physics, quantum computers are physical devices manipulated according to the laws of quantum physics.

Quantum Computers and the 20 Questions Game

Having understood that a computation is ultimately a physical process, let's go on to see how using a quantum machine is much like playing the game of 20 questions. Twenty questions is played as follows. I think of a number between 1 and 2^{20} . You try to guess my secret number by asking questions such as, "Is your secret number less than 2378?" If you ask your questions well, you can guess my secret number in, at the most, 20 questions. Why? Well, with each question, you can eliminate half of the remaining candidates. Computer scientists call this process binary search, and it allows you to find a secret number less than 2^n in $\log 2^n = n$ questions at the most. The key idea is that, by cutting the number of possibilities in half with each question, you are left with one possibility after only n questions. This principle generalizes. For example, if you are searching for a secret item among N possibilities and with each question you are able to eliminate a fraction $1 - 1/c$ of the possibilities, then you can find the secret in $\log_c N$ questions. In general, you might not be looking for a number. You might be looking for a secret element x in a set S called a search space. The key to quick success is still to be able to eliminate a constant fraction of the remaining candidates. Now, let's consider a slightly different version of this game, which we call "random 20 questions."

In playing random 20 questions, you don't get to choose your question. Instead, you randomly select a subset Q (used for the word "question") consisting of half of the N elements in the search space, and you ask, "Is the secret element in Q ?" After I give you the honest answer, you choose a new random subset Q and ask again. Surprisingly, again after only about $\log N$ questions, you will almost surely have narrowed the possibilities down to the one correct answer. We say "almost surely" because there is a tiny, tiny chance that you will get unlucky and never be able to eliminate one of the elements that is not the secret element. This tiny chance is the result of each question having been selected randomly rather than deterministically, which is the case when playing the original 20 questions game. After $2 \log N$ questions, for example, that possibility is incredibly small. So, even by asking random questions, you can discover the secret element quickly. The reason is that, as in the original 20 questions game, you are able to eliminate each incorrect element as a possibility. Although in the random 20 questions game this process of elimination is only very highly probable, it is so close to being certain that, for all practical purposes, we won't worry about it. Now, let's talk about playing quantum 20 questions.

In this game, I choose a secret quantum state ρ_1 from a search space of quantum states $S = \{\rho_1, \rho_2, \dots, \rho_N\}$, and I supply a copy of the secret state whenever you request one. Your task is to discover my secret quantum state by asking quantum questions, that is, by doing measurements on each requested quantum state and thus getting information about the state. Now, let's back up a bit and clarify these terms. What is a quantum state? A pure state ψ is simply a vector in a Hilbert space. A mixed state, or more simply a state, is a convex combination of pure states ψ_i , that is, a classical probabilistic mixture of pure states:

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i|, \quad \text{where } \sum_i p_i = 1. \quad (1)$$

Quantum Questions

What is a quantum question? A quantum question is typically called an observable. We'll think of a quantum question as simply an orthonormal basis. The answer to a quantum question will be one of the basis vectors. So, suppose the secret quantum state is a pure state $|\phi\rangle$ and the quantum question is $\{|\phi_1\rangle, |\phi_2\rangle, \dots, |\phi_M\rangle\}$, a basis of the M -dimensional Hilbert space. According to the basic rules of quantum mechanics, we get the answer $|\phi\rangle$ with probability $|\langle\phi|\phi\rangle|^2$. If we have a mixed state instead of a pure state, the probability formula is extended by convexity, as usual. How many quantum questions does it take to guess the secret quantum state? That depends on lots of things. It depends on what quantum questions you are allowed to ask me. And it also depends on how different the states in S are from each other. In this context, the word "different" means how distinguishable the states are from each other. For example, two orthogonal pure states are as different as two states can be. Two very nearly parallel pure states are almost indistinguishable in that it takes many experiments and questions to tell them apart based on the outcome statistics. The standard measure of similarity between two pure states is simply their overlap $\langle\phi|\phi\rangle$. There are measures for the similarity or overlap of mixed states as well, but we won't need the formula. We just need to know that to tell apart two similar states requires many experiments whereas to tell apart two very different states requires few experiments.

So, going back to quantum 20 questions, let's assume you can ask any quantum question you want; that is, you can choose any orthonormal basis as the observable. If all the states in S are sufficiently different from each other, you can find my secret state after only a few questions. Usually, when we use the word "few" in this context, we mean $\log |S|$ or $\log^2 |S|$ or something like that. (A computer scientist would say that few means a polynomial function of the logarithm of the size of the search space.) The key to a fast search is that all the states must be quite different from each other.

It turns out that playing search games is much like trying to break codes. If you try to break a code, you want to look for a cryptographic key. The key is what allows you to decipher the code and read the message. One popular code is the RSA. Named after its inventors—Ronald Rivest, Adi Shamir, and Leonard Adleman—the RSA uses as its key the secret factors of a large number N . Now, suppose you are trying to break a code by finding a secret key k from among a very large set of possible keys $K = \{k_1, k_2, \dots, k_M\}$. Further suppose that, by some process and without knowing the key, you can prepare a quantum state ρ corresponding to the key k . So, you now have a state ρ , which you know comes from the search space $S = \{\rho_1, \dots, \rho_M\}$, which is the set of states corresponding to all the possible secret keys, but you don't know exactly which of the states you have. If the states of S are all sufficiently different, then you can ask quantum questions to determine the secret state efficiently. And if you can find the secret state, then you can easily figure out the original secret key corresponding to that secret state!

Indeed, this is precisely how quantum computers would solve various classical cryptographic problems, such as factoring and finding discrete logarithms. A factoring problem is one in which you are given a very large number N (say, one with 2000 digits), which is the product of two primes $N = pq$, and your task is to find p and q . For the discrete logarithm problem, you are given a large prime number p (say, once again, one with 2000 digits) and two numbers a and b less than p . Your task is to find n such that $a^n = b \pmod{p}$. In both cases, you are looking for a secret key k from among a known set of possible secret keys. Also, in both cases there is a process by which you can prepare a quantum state from which k can be deduced. Significantly, this preparation process does not require knowing k .

This last point is important because, if you had to know the key first, then the code-breaking machine would not be very useful. We will later illustrate this process in an

Playing search games is much like trying to break codes. If you try to break a code, you want to look for a cryptographic key. To solve classical cryptographic problems with quantum computers, you are looking for a secret key from among a known set of possible secret keys.

example (see the section “Simon’s Problem”). Finally, this process has the special and important quality that, for two different keys, k_1 and k_2 , the resulting quantum states, ρ_1 and ρ_2 , are quite different, or clearly distinguishable from one another, as discussed before. We can therefore ask quantum questions, which allow us to distinguish among states and identify secret keys. This ability to distinguish among the states is usually accomplished by eliminating the possibility of a constant fraction, say $1/2$, of the remaining states. As we saw in the game of 20 questions, eliminating a constant fraction after each question allows us to narrow the possible states down to the one true state in only $\log N$ questions. However, since the quantum formula gives probabilities for certain outcomes, we eliminate the false states with high probability (not with certainty), as in the game of random 20 questions.

Identifying Secret Quantum States

Let us fill in some of the technical details of our sketch. First, can we really ask any quantum question? No, we can’t, but fortunately we are able to ask the questions that let us solve factoring and discrete logarithm problems. Recalling our observation that a computation is actually a physical process, we must be sure to carry out efficiently the physical process corresponding to the quantum question we wish to ask. We accomplish this task by breaking down the observable into elementary quantum “gates.” Elementary quantum gates are analogous to the basic logical gates **and**, **or**, and **not**, which are the building blocks of circuits in classical computers (for more details, see Shor 1997). In the case of factoring and discrete logarithm problems, it turns out that we have to ask only one quantum question over and over again in order to obtain enough information for identifying the secret quantum state. Called the quantum Fourier transform, this quantum question allows us to distinguish among the states that arise in the two search spaces for the factoring and discrete logarithm problems. These states are called hidden subgroup states because, in those problems, the key we are looking for corresponds to an unknown subgroup H of a finite abelian group G . The search space corresponds to the set $\{\rho_{H_1}, \rho_{H_2}, \dots, \rho_{H_i}\}$, where H_1 to H_i is a range over all the possible subgroups of G , and ρ_H is the mixed state that corresponds to a uniform mixture of the pure coset states

In factoring and discrete logarithm problems, we must ask only one quantum question over and over again in order to obtain information for identifying the secret quantum state.

$$|c + H\rangle = \frac{1}{\sqrt{|H|}} \sum_{h \in H} |c + h\rangle . \quad (2)$$

It can be shown that for H_1 and H_2 , two different subgroups, the corresponding states ρ_{H_1} and ρ_{H_2} are sufficiently different. Mathematically speaking, the overlap of ρ_{H_1} and ρ_{H_2} is less than $1/2$ (Ettinger et al. 1999). For a discussion of the hidden subgroup problem and the reasons why the quantum Fourier transform is the right quantum question, see Ettinger and Peter Hoyer (1999).

Simon’s Problem

To illustrate everything we have discussed, let’s consider a concrete example known as Simon’s problem. Simon’s problem and the quantum algorithm to solve it contain the essence of what is going on in the factoring and discrete logarithm problems; the latter set of problems, however, also contains a number of technical twists that obscure the main ideas. The set of all bit strings of length n , denoted Z_2^n , is a commutative group if

Our quantum algorithm for solving Simon's problem allows distinguishing among different states and thus discovering the underlying secret bit string.

we add bit strings using “binary add without carry.” This group will be our search space. I will secretly choose an element s of this group and provide you with a function in the form of a “black box,” f_s on Z_2^n , with the following special property: I guarantee that $f_s(x) = f_s(y)$ if and only if $x - y = s$. So, the function f_s encodes the secret bit string s . Because f depends entirely on s , the latter becomes a subscript on f . If you compute the function on the elements of the group $f_s(a), f_s(b), f_s(c), \dots$, eventually you'll get a collision, which means that you'll find $f_s(g) = f_s(t)$ and then you'll know that the secret bit string is $s = g - t$. But notice that the search space, or the group, has 2^n elements, which is a very large number. In the worst case, it could take you $2^{n-1} + 1$ calculations to get a collision, and on average it will take about $2^{n/2}$ because of the so-called birthday paradox.¹ That is still a lot of time! But the quantum algorithm can solve this problem much more quickly—in about n tries only.

Here is how Simon's problem works: You start with a quantum computer whose qubits are conceptually divided into two registers. Then you prepare the pure state $|\psi\rangle = 1/2^{n/2} \sum_b |b\rangle$, where $b \in Z_2^n$. Thus, in the first register, there is a superposition of all the b s. Now, because you have the black box function f_s , you can compute $f_s(b)$ in the second register to obtain the pure state $|\psi_s\rangle = 1/2^{n/2} \sum_b |b\rangle |f_s(b)\rangle$, where again $b \in Z_2^n$. Notice that this procedure for preparing the state ψ_s is easily accomplished without any knowledge of the secret bit string s . Of course, for different secret bit strings, we obtain different states. In fact, this is the key point. Our quantum algorithm is really just a method used to distinguish among these different states and thus discover the underlying secret bit string.

We now observe, or perform a measurement, on the second register. Because of the way quantum mechanics works, this observation collapses $|\psi_s\rangle$, producing a specific value in the second register, say c , and the first register is left in a superposition of bit strings that map to c under f_s . Because f_s has the special property described earlier, the bit strings that map to c will differ by the secret bit string s . Therefore, the state of the computer is

$$|\psi\rangle_{a,s} = \frac{1}{\sqrt{2}} |a\rangle |c\rangle + \frac{1}{\sqrt{2}} |a+s\rangle |c\rangle, \quad (3)$$

where a and $a+s$ are elements of Z_2^n such that $f_s(a) = c$ and $f_s(a+s) = c$. The only use of the second register is to produce this special superposition in the first register. We will no longer use the second register or its contents, so we drop it from our notation and write

$$|\psi\rangle_{a,s} = \frac{1}{\sqrt{2}} |a\rangle + \frac{1}{\sqrt{2}} |a+s\rangle. \quad (4)$$

When c is chosen, the resulting mixed state can be written as

$$\rho_s = \frac{1}{2^{n-1}} \sum_a |\psi\rangle_{a,s} \langle\psi|_{a,s}. \quad (5)$$

Recall that we don't know the secret bit string s , and therefore we don't know that the state we just prepared is ρ_s . All we know is that we have prepared a state that is in the search space of quantum states $\{\rho_s\}_{s \in Z_2^n}$. Each of these possible quantum states corresponds to a possible secret bit string. Our task is to identify the secret quantum state

¹ The birthday paradox derives its name from the surprising result that you only need 23 people (a slightly larger number than $365^{1/2}$) to have a 50 percent chance that at least two of them have the same birthday.

and thus the secret bit string. We now define the Fourier observable. For each bit string b in Z_2^n , define

$$|\chi_b\rangle = \frac{1}{\sqrt{2^n}} \sum_{d \in Z_2^n} (-1)^{b \cdot d} |d\rangle, \quad \text{where } b \cdot d = \sum_i b_i d_i \pmod{2}. \quad (6)$$

The orthonormal basis is $\{|\chi_b\rangle\}$, where $b \in Z_2^n$ is called the Fourier basis or the Fourier observable. Mathematicians might recognize this basis as being composed of the characters of the group Z_2^n . A character χ of a finite abelian group is a homomorphism from the group to the circle in the complex plane. Formally, the Hilbert space in which we are working is $C[G]$, the group algebra, which is the complex vector space with the canonical basis, or the point mass basis, indexed by the elements of the group. A character can be viewed as a vector in $C[G]$ via the following identification:

$$|\chi\rangle = \frac{1}{\sqrt{|G|}} \sum_{g \in G} \chi(g) |g\rangle. \quad (7)$$

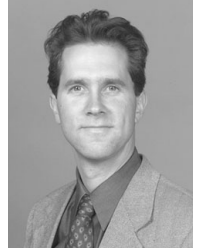
It is a fundamental fact (Tolimieri et al. 1997) that the set of characters viewed as vectors in this way is an orthonormal basis for $C[G]$. Indeed, a Fourier transform is nothing other than a change of basis from the point mass basis, $\{|g\rangle\}_{g \in G}$, to the basis of characters, $\{|\chi\rangle\}_\chi$.

It is easy to show (Jozsa 1998) that, if we now observe the contents of the remaining register in the Fourier basis, we observe $|\chi_b\rangle$ with nonzero probability if and only if $s \cdot b = 0 \pmod{2}$. This is the important relationship between the secret bit string s and the only possible outcomes of the experiment. Therefore, if the actual outcome of the observation is $|\chi_b\rangle$, then we have eliminated half of the possible secret states. We have therefore eliminated all states ρ_d such that $d \cdot b = 1 \pmod{2}$. By repeating the state preparation procedure followed by a measurement in the Fourier basis approximately n times, we eliminate all possible states except the true secret state ρ_s . ■

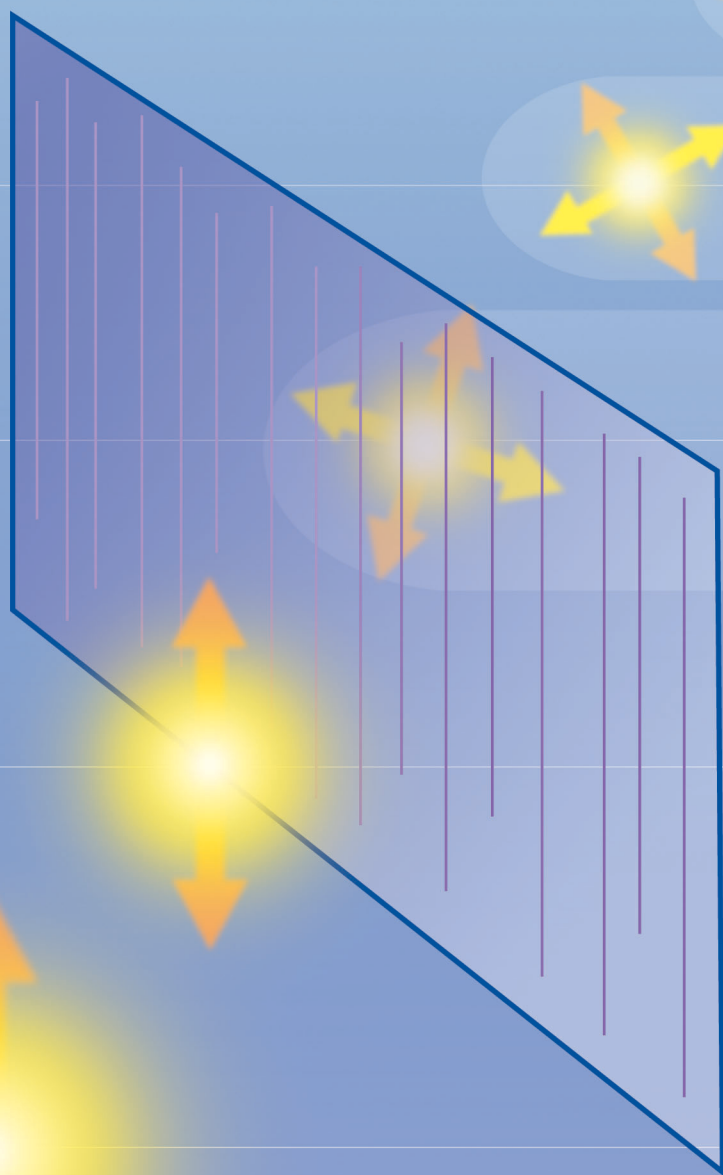
Further Reading

- | | |
|---|---|
| <p>Ettinger, M., and P. Hoyer. 1999. Quantum State Detection via Elimination [Online]: http://eprints.lanl.gov/quant-ph/9905099.</p> <p>Ettinger, M., P. Hoyer, and E. Knill. 1999. Hidden Subgroup States are Almost Orthogonal. [Online]: http://eprints.lanl.gov/quant-ph/9901034.</p> <p>Jozsa, R. 1998. Quantum Algorithms and the Fourier Transform. <i>Proc. R. Soc. London, Ser. A</i> 454: 323.</p> | <p>Shor, P. W. 1997. Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on Logarithms on a Quantum Computer. <i>SIAM J. Computing</i> 26: 1484.</p> <p>Tolimieri, R., M. An, and C. Lu. 1997. <i>Mathematics of Multidimensional Fourier Transform Algorithms</i>. New York: Springer.</p> |
|---|---|

Mark Ettinger graduated from the Massachusetts Institute of Technology in 1987 with Bachelor's degrees in physics and mathematics. In 1996, Mark received his Ph.D. in mathematics from the University of Wisconsin at Madison.



He first came to Los Alamos National Laboratory in 1993, as a graduate student, then entered the postdoctoral program after graduation in 1996, and became a staff member in 1999. Mark worked on the group-theoretical approach to quantum algorithms for four years and is now primarily interested in (classical) algorithmic problems in postgenomic computational biology.



“When two systems, of which we know the states by their respective representatives, enter into temporary physical interaction due to known forces between them, and when after a time of mutual influence the systems separate again, then they can no longer be described in the same way as before, viz. by endowing each of them with a representative of its own. I would not call that one but rather the characteristic trait of quantum mechanics, the one that enforces its entire departure from classical lines of thought. By the interaction, the two representatives (or ψ -functions) have become entangled.”

—Erwin Schrödinger (1935)

Quantum State Entanglement

Creation, characterization, and application

Daniel F. V. James and Paul G. Kwiat

Entanglement, a strong and inherently nonclassical correlation between two or more distinct physical systems, was described by Erwin Schrödinger, a pioneer of quantum theory, as “the characteristic trait of quantum mechanics.” For many years, entangled states were relegated to being the subject of philosophical arguments or were used only in experiments aimed at investigating the fundamental foundations of physics.

In the past decade, however, entangled states have become a central resource in the emerging field of quantum information science, which can be roughly defined as the application of quantum physics phenomena to the storage, communication, and processing of information.

The direct application of entangled states to quantum-based technologies, such as quantum state teleportation or quantum cryptography, is being investigated at Los Alamos National Laboratory, as well as other institutions in the United States and abroad. These new technologies offer exciting prospects for commercial applications and may have important national-security implications. Furthermore, entanglement is a *sine qua non* for the more ambi-

tious technological goal of practical quantum computation.

In this article, we will describe what entanglement is, how we have created entangled quantum states of photon pairs, how entanglement can be measured, and some of its applications to quantum technologies.

Classical Correlation and Quantum State Entanglement

To describe the concept of quantum entanglement, we are first going to describe what it is not! Let us imagine the simple experiment illustrated in Figure 1. In that experiment, a source S_1 continually emits pairs of photons in two directions. As seen in the figure, one photon goes toward an observer named Alice, while the other goes toward Bob.

First, imagine that the photons emitted by S_1 are always polarized in the horizontal direction. Mathematically, we say that each photon is in the pure state denoted by the ket $|H\rangle$, that is, the “representative” of the state Schrödinger referred to in the quotation on the opposite page. Because the photons are paired, the combined state of the

two photons is denoted $|HH\rangle$, where the first letter refers to Alice’s photon and the second to Bob’s.

Alice and Bob want to measure the polarization state of their respective photons. To do so, each uses a rotatable, linear polarizer, a device that has an intrinsic transmission axis for photons. For a given angle ϕ between the photon’s polarization vector and the polarizer’s transmission axis, the photon will be transmitted with a probability equal to $\cos^2\phi$. (See the box “Photons, Polarizers, and Projection” on [page 76](#).) Formally, the polarizer acts like a quantum-mechanical projection operator P_ϕ selecting out the component of the photon wave function that lines up with the transmission axis. We say that the polarizer “collapses” the photon wave function to a definite state of polarization. If, for example, the polarizer is set to an angle θ with respect to the horizontal, then a horizontally polarized photon is either projected into the state $|\theta\rangle$ with probability $\cos^2\theta$ or absorbed with probability $1 - \cos^2\theta = \sin^2\theta$. The bizarre aspect of quantum mechanics is that the projection process is probabilistic. The fate of any given photon is completely

This article is dedicated to the memory of Professor Leonard Mandel, one of the pioneers of experimental quantum optics, whose profound scientific insights and gentlemanly bearing will be sorely missed.

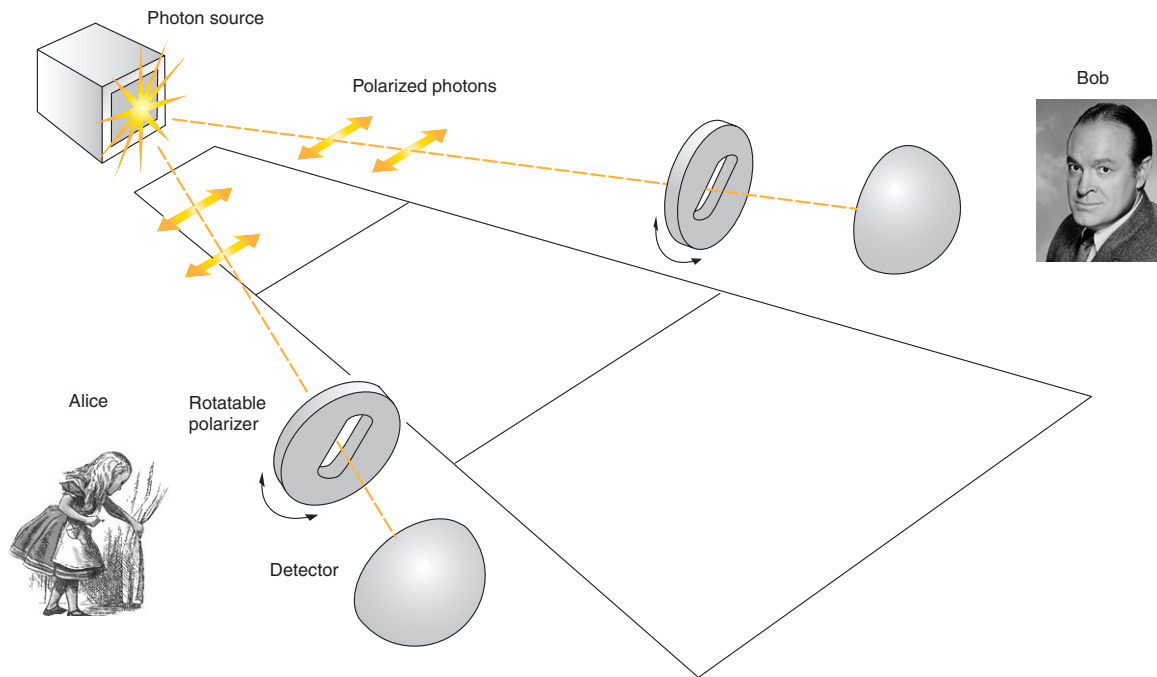


Figure 1. A Simple Two-Photon Correlation Experiment

In this experiment, a source emits pairs of photons: One photon is going to Alice and the other to Bob. Each photon passes through a linear polarizer on its way to its respective detector. Both Alice and Bob's polarizers are rotatable and can be aligned to any angle with respect to the horizontal, but Bob's is always kept parallel to Alice's. For a given polarizer setting, Alice and Bob record those instances when they have the same results, that is, when both detect photons or when they don't. The figure shows the source emitting two horizontal photons in the state $|\Psi\rangle = |HH\rangle$. The experiment can be performed with other sources to examine differences between other two-photon states. (Picture of Bob is courtesy of Hope Enterprises, Inc.)

unknown. Furthermore, any information about the photon's previous polarization state is lost.

Getting back to the experiment, we assume that Alice and Bob's polarizers are always aligned in the same way: When Alice sets her polarizer to a certain angle, she communicates her choice to Bob, who uses the same setting. Behind each polarizer is a detector. In our experiment, Alice and Bob rotate their polarizers to a certain angle with respect to the horizontal and record whether they detect a photon. Then, they repeat the procedure for different polarizer settings. If Alice looks only at her own data (or Bob looks only at his), she can determine the polarization state of the photons emitted by the source—see Figure 2(a). But Alice and Bob can also make a photon-per-photon comparison of their data and determine the probability that they have the

same result, that is, they can examine the photon–photon correlations.

Suppose Alice has her polarizer oriented to transmit horizontally polarized photons. In that case, each photon coming to her from S_1 will be transmitted, and her detector will “click,” indicating a photon has arrived. Subsequent communication with Bob would reveal that he also detected each photon, so at this polarizer setting, there is a perfect correlation between Alice's detection of a photon and Bob's. Similarly, by rotating the polarizer to the vertical position, the two would again discover a perfect correlation, namely, neither party would detect his or her photons.

The correlation changes when Alice and Bob have their polarizers oriented, say, at $+45^\circ$ to the horizontal. In that case, the photon sent to Alice has a 50 percent chance of passing through her polarizer, and independently, the

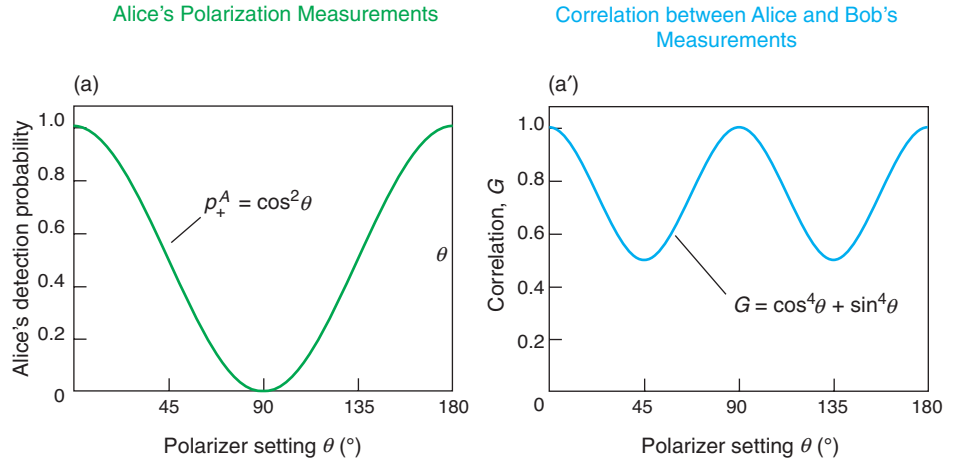
photon sent to Bob has a 50 percent chance of passing through his.

The probability is therefore 25 percent that both Alice and Bob detect a photon, 25 percent that neither detects a photon, and thus 50 percent that they obtain the same result.

The correlation function G is shown in Figure 2(a'). It is equal to the product of the independent probabilities for detecting a photon $[(\cos^2\theta)_A \times (\cos^2\theta)_B]$, plus the product of the probabilities for not detecting one $[(\sin^2\theta)_A \times (\sin^2\theta)_B]$, where subscripts A and B are for Alice and Bob, respectively. Thus, Alice and Bob deduce that the two photons are independent of each other and the wave function is in fact separable: $|HH\rangle = |H\rangle|H\rangle$. In other words, the correlation is entirely consistent with classical probability theory. The photons are classically correlated.

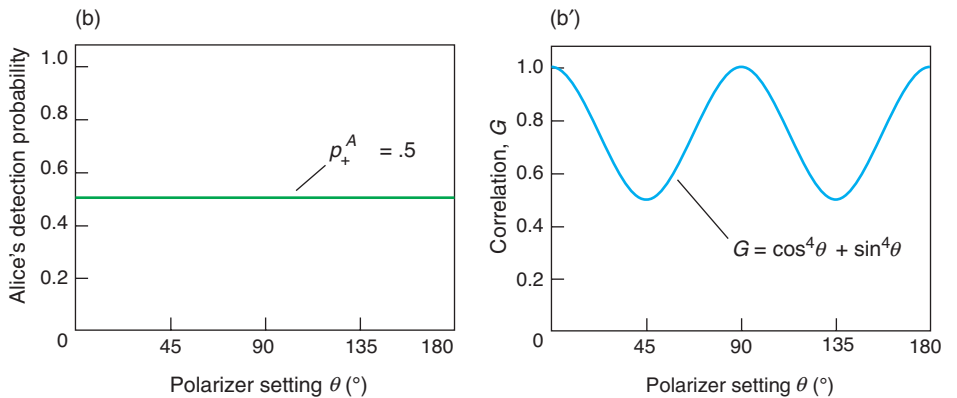
Now, consider performing the

(a) S_1 emits photons in the pure state $|HH\rangle$. Alice measures a $\cos^2\theta$ function for her polarization data and deduces that photons coming to her are horizontally polarized. (A different linear polarization would shift the curve to the left or right.) (a') We define the correlation function G as the probability that both Alice and Bob detect a photon, plus the probability that neither detects a photon. For this source, G is completely consistent with classical probability theory for independent events; that is, the correlation function is the product of the detection probability of each photon in the pair.



Probability that Alice (or Bob) detects a photon: $p_+ = \cos^2\theta$.
 Probability that Alice (or Bob) does not detect a photon: $p_- = \sin^2\theta$.
 For independent photons: $G = G_{HH} = p_+^A \times p_+^B + p_-^A \times p_-^B = \cos^4\theta + \sin^4\theta$.

(b) The source S_2 emits photons in the partially mixed state $1/2(|HH\rangle\langle HH| + |VV\rangle\langle VV|)$. Photons from this source do not have a net polarization. Alice receives at random either an $|H\rangle$ or a $|V\rangle$ photon, so the sum of her measurements averages to a 50 percent detection probability independent of angle. (b') The correlation function G , however, is the same as in (a), revealing that the photons in each pair are independent of each other and have polarization H or V. Therefore, the two photons exhibit the same classical correlations seen in (a).



For this mixed state,
 $G = 1/2(G_{HH} + G_{VV}) = G_{HH}$.

(c) The source S_3 emits photons in the maximally entangled state $1/\sqrt{2}(|HH\rangle + |VV\rangle)$. Unlike the photons in the mixed state, each photon is unpolarized. Nevertheless, if Alice and Bob align their polarizers the same way, they always get the same result independent of angle. (c') Polarization measurements of the two photons are 100 percent correlated. The photons exhibit "quantum" correlations.

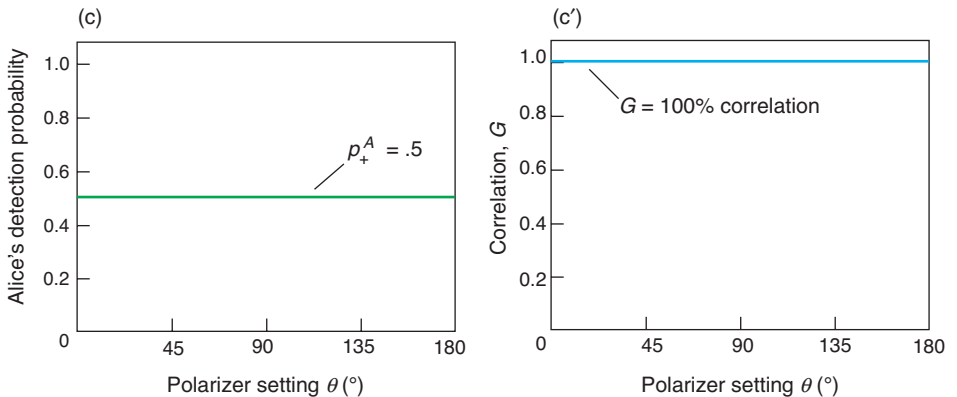


Figure 2. Quantum States, Polarization, and Correlation

The three sets of graphs show the results of the three experiments discussed in the text. In each case, the leftmost graph shows the probability that Alice alone detects a photon and reveals information about the net polarization state of her photon. The rightmost graph shows the probability that Alice and Bob have the same result, which reveals information about the two-photon state.

Pure, Entangled, or Mixed?

A pure state is a vector in a system's Hilbert space. For example, the most general, pure two-photon polarization state can be written as

$$|\psi_{\text{pure}}\rangle = \alpha|HH\rangle + \beta|HV\rangle + \gamma|VH\rangle + \delta|VV\rangle . \quad (1)$$

This state is specified by the four probability amplitudes α , β , γ , and δ (expressed by four complex numbers or eight real numbers) although these parameters are subject to two constraints. The first is that the mean-square amplitudes must equal unity, that is,

$$|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1 . \quad (2)$$

The second relates to the fact that the overall phase of a wave function has no physical relevance. The net result of these constraints is that any pure two-photon state depends on only six independent real numbers.

In general, however, any physical system contains a greater or lesser degree of randomness and disorder, and one must adapt the formalism of quantum mechanics to take this randomness into account. We do so by averaging over the fluctuations. It is convenient to represent states as density operators, or density matrices, formally defined as

$$\rho = \overline{|\psi\rangle\langle\psi|} , \quad (3)$$

where the overbar denotes an ensemble average over the randomness. All the measurable properties of the state are determined by ρ .

The density matrix must be used when representing mixed states, which can be thought of as probabilistic combina-

tions of pure states. Mathematically, the density matrix can always be decomposed into an incoherent sum over pure states,

$$\rho = \sum_i p_i |\psi_i\rangle\langle\psi_i| , \quad (4)$$

where each $|\psi_i\rangle$ is a pure state and p_i are probabilities with values that lie between 0 and 1 and whose sum is 1. In general, this decomposition is not unique. To characterize mixed states, one uses mean values and classical coherences; that is, one must specify the four mean-square amplitudes (subject to the constraint $|\alpha|^2 + |\beta|^2 + |\gamma|^2 + |\delta|^2 = 1$) and the six independent classical complex correlations $\overline{\alpha^* \beta}$, $\overline{\alpha^* \gamma}$, and so on.

For example, the source S_2 mentioned in the text emits a partially mixed state that is 50 percent $|HH\rangle$ and 50 percent $|VV\rangle$, so that

$$\rho_{\text{mix}} = 0.5 |HH\rangle\langle HH| + 0.5 |VV\rangle\langle VV| , \quad (5)$$

or in matrix form

$$\rho_{\text{mix}} = \begin{bmatrix} .5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & .5 \end{bmatrix} . \quad (6)$$

This state is neither pure nor completely random; it is partially mixed.

We next consider whether quantum states involving two or more systems (for example, two photons), are separable

experiment with a second source S_2 that has a 50 percent chance to emit two horizontally polarized photons $|HH\rangle$ and a 50 percent chance to emit two vertically polarized photons $|VV\rangle$. This type of source emits photons in a mixed state, which cannot be written as a single “ket.” Instead, a mixed state must be analyzed in terms of several kets, each representing a particular, distinct pure state that has a probability associated with it. Making a measurement on a mixed state is

equivalent to probing an ensemble of pure states. The likelihood of measuring a particular pure state is given by the appropriate probability. (More-detailed, mathematical descriptions of pure and mixed quantum states are found in the box “Pure, Entangled, or Mixed?” above.)

The output of S_2 is random (either $|HH\rangle$ or $|VV\rangle$), so Alice receives at random either an $|H\rangle$ or a $|V\rangle$ photon. Because the probability of detecting $|H\rangle$ is $1/2 \cos^2\theta$, and the probability of

detecting $|V\rangle$ is $1/2 \sin^2\theta$, Alice has a 50 percent chance of detecting a photon regardless of how she sets her polarizer. The same is true for Bob. Each observer, therefore, deduces that the photons coming from S_2 have no net polarization. But as seen in Figure 2(b'), the correlation function tells a different story. In fact, the correlation function for this source is identical to the one obtained for S_1 because, in both cases, the individual photons leave the source in definite

or entangled. If the state is separable and pure, it can be written (in some basis) as a product of the states of the individual systems, that is, as

$$|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle, \quad (7)$$

where \otimes denotes the tensor product. The state $|\psi_1\rangle = |HH\rangle$ is one such product of pure states and can be written as

$$|\psi_1\rangle = |H_A\rangle \otimes |H_B\rangle. \quad (8)$$

Another example is the state

$$|\psi\rangle = (|HH\rangle + |HV\rangle + |VH\rangle + |VV\rangle)/2, \quad (9)$$

which can be written as the product state

$$|\psi\rangle = 1/\sqrt{2}(|H\rangle + |V\rangle)_A \otimes 1/\sqrt{2}(|H\rangle + |V\rangle)_B. \quad (10)$$

A third example is the matrix ρ_{mix} on the opposite page, which represents a separable mixed state.

In contrast, if there is no way to write the two-photon state as a direct product of states, the state is said to be entangled. This definition leads to a quantity called concurrence, which is defined for the general pure state $|\psi_{\text{pure}}\rangle$ by

$$C = 2|\alpha\delta - \beta\gamma|. \quad (11)$$

If and only if C is zero is the state separable. If C is equal to unity (its maximum value), the state is maximally entangled.

For example, consider any one of the four Bell states

$$|\Phi_{\pm}\rangle = 1/\sqrt{2}(|HH\rangle \pm |VV\rangle), \text{ and}$$

$$|\Psi_{\pm}\rangle = 1/\sqrt{2}(|HV\rangle \pm |VH\rangle). \quad (12)$$

These states are a basis for the two-photon Hilbert space, and linear combinations of the four states can be used to represent any two-photon state. If we compare, say, $|\Phi_{+}\rangle$ with the general state $|\psi_{\text{pure}}\rangle$, we have $\alpha = \delta = 1/\sqrt{2}$, and $\beta = \gamma = 0$. Thus $C = 1$, and this Bell state is maximally entangled (as are the other three).

The value of C provides a good metric for the amount of entanglement in a pure two-qubit system. Equivalently, some researchers use C^2 (a quantity known as the tangle) to characterize the degree of entanglement.

The concurrence can also be defined for mixed states, although the definition is much more complicated. Indeed, calculating the concurrence for mixed states of more than two qubits is currently a hot topic of research.

In the everyday world, it is common to ascribe two (or more) variables to the same object (for example, a hot, sweet cup of coffee). Similarly, quantum states are described by the two characteristics discussed above, so that it is possible to have a pure entangled state, a pure separable state, a mixed separable state, or something in between, such as a partially mixed, partially entangled state.

polarization states. For S_1 , the polarization information is “carried” individually by each photon. For S_2 , the polarization information is carried by the photon pairs. By examining the correlations, Alice and Bob can deduce that information.

A different situation occurs for a source S_3 that emits pairs of photons in the state $|\Phi_{+}\rangle = 1/\sqrt{2}(|HH\rangle + |VV\rangle)$. Like the mixed state from S_2 , this state is a combination of two horizontally polarized photons and two verti-

cally polarized photons. Unlike the mixed state, $|\Phi_{+}\rangle$ is a coherent, quantum mechanical superposition: A probability amplitude is associated with each component, $|HH\rangle$ and $|VV\rangle$, and the two components have a fixed phase relationship. An important property of this particular state is that we can rotate the axes of polarization (H and V) and not change the state’s essential properties.

The state $|\Phi_{+}\rangle$ is a fully entangled quantum state. It cannot be factorized,

or separated, into a part describing one of the photons and a part describing the other. The two photons are inextricably linked to each other and their properties are always correlated.

A measurement of one of the photons makes the two-photon state instantly disappear, and the remaining photon assumes a definite state that is perfectly correlated with the measured photon. Neither photon carries definite information by itself—all the information is carried in the joint two-photon state.

Thus, when Alice and Bob repeat the experiment using the source S_3 , the correlation is 100 percent regardless of polarizer orientation (assuming Bob's polarizer is always set the same way as Alice's). Figure 2(c) illustrates the striking difference between the classical correlations of the photons generated by the sources S_1 and S_2 and the nonclassical correlations exhibited by entangled photons.

To better understand the correlation curve shown for $|\Phi_+\rangle$, consider that quantum mechanics allows us to express that state in any basis; that is, $|\Phi_+\rangle = 1/\sqrt{2} (|XX\rangle + |YY\rangle)$, where $|X\rangle$ is an arbitrary linear-basis state and $|Y\rangle$ is the orthogonal-basis state. Suppose Alice has her polarizer set to $+45^\circ$. In the diagonal $(+45/-45)$ basis, the entangled state will be $|\Phi_+\rangle = 1/\sqrt{2} (|+45,+45\rangle + |-45,-45\rangle)$. If Alice detects her photon (a 50–50 proposition), then Bob's photon will collapse to the $|+45\rangle$ state, and he will detect his photon as well. Likewise, if Alice doesn't detect her photon, Bob won't detect his. The same deductions can be made for any polarizer setting.

According to quantum mechanics, the correlation occurs regardless of the distance separating the two photons. For example, suppose one of two entangled photons from the state $|\Phi_+\rangle$ is sent to Alice, who "stores" it in her laboratory at Los Alamos, New Mexico. The other photon is sent to Bob, who is in orbit about the star α -Centauri, nearly 4 light-years away. After some time, Alice performs a measurement on her photon and determines that it is $|H\rangle$. Her measurement selects the $|HH\rangle$ part of the state $|\Phi_+\rangle$ and eliminates the $|VV\rangle$ part so that Bob's photon is necessarily in the state $|H\rangle$. If, instead, Alice has determined that her photon was $|+45\rangle$, the state of Bob's photon will be instantly collapsed to $|+45\rangle$ as well. In other words, the state of Bob's photon has been nonlocally influenced by Alice's measurement. By nonlocal, we

mean that the correlation between Alice and Bob's measurements occurs even if there is not enough time for a light signal (or any signal) to propagate between the two experimentalists. This is not to say that special relativity has been violated: Because Alice cannot predetermine the outcome of her measurement, she cannot use the nonlocal quantum correlations to send any information to Bob. In fact, entanglement can never be used to send signals faster than the speed of light. Nonetheless, Bob's photon "knows" the outcome of Alice's measurement.

Nonlocality was the central point of a famous argument raised by Albert Einstein, Boris Podolsky, and Nathan Rosen in 1935, now known as the EPR paradox. The three physicists disagreed with the Copenhagen interpretation of quantum mechanics, according to which the state of a quantum system is indeterminate until it is projected into a definite state as a result of a measurement. Einstein, Podolsky, and Rosen argued that even unmeasured quantities corresponded to definite "elements of reality." The quantum state only appeared to be indeterminate because some of the parameters that characterize the system were unknown and unmeasurable. These local parameters, or "hidden variables," determined the outcome of the experiment.

In 1964, John Bell showed that the correlations between measured properties of any classical two-particle system would obey a mathematical inequality, but the same measured correlations would violate the inequality if the two particles were an entangled quantum system. Experiments could therefore determine if nature exhibited nonlocal features. Following the

development of laboratory sources of entangled photons, experimental tests of Bell's inequality were pursued with vigor. The results to date suggest that the observed photon correlations cannot be explained by any local hidden-variable theory,¹ and most physicists agree that quantum mechanics is truly a nonlocal theory.

Entanglement and Quantum Information

Entanglement, a measurable property of quantum systems, can be exploited for specific goals. Here, we present three potential applications, all of which have been shown to work as proof-of-principle demonstrations in the laboratory.

Quantum Cryptography. Consider two bank managers, Alice and Bob, who want to have a secret conversation. If they are together in the same room, they can simply whisper discretely to each other, but when Alice and Bob are in their respective cross-town offices, their best chance for secret communication is to encrypt their messages.

A generic, classical encryption protocol would begin when Alice and Bob convert their messages to separate binary streams of 0s and 1s. Encryption (locking up the messages) and decryption (unlocking the messages) are then performed with a set of secret "keys" known only to the two bankers. Each key is a random string of 0s and 1s that is as long as the binary string comprising each message. To encrypt, Alice (the sender) sequentially adds each bit of

¹ There were two loopholes to the EPR tests. The first stemmed from the fact that the detectors were not efficient enough. Consequently, the observed correlations could have been the result of some new physics that did not require nonlocal interactions. The second loophole stemmed from the researchers' inability to choose rapidly and randomly a basis for photon measurement. This inability allowed for a potential communication conspiracy between Alice and Bob's systems. Both of these loopholes have recently been closed but, so far, not in the same experiment.

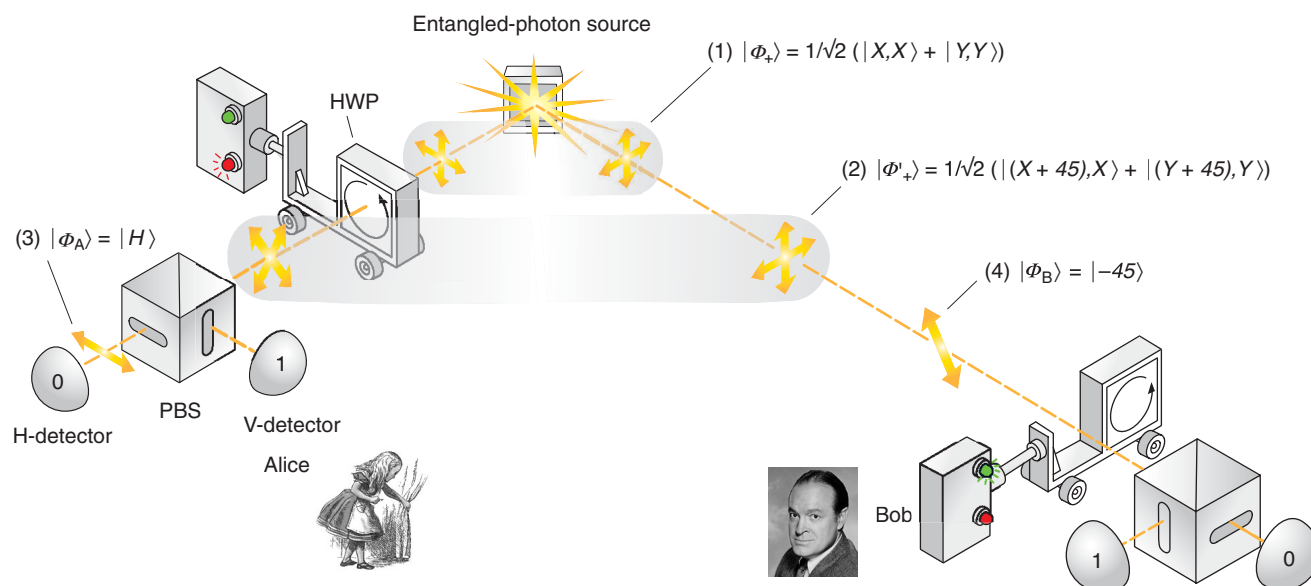


Figure 3. Quantum Cryptography Using Entangled Photons

Alice and Bob can use the properties of entangled photons to create a pair of identical cryptographic keys. (1) The source emits entangled photons in a maximally entangled state $|\Phi\rangle = 1/\sqrt{2}(|XX\rangle + |YY\rangle)$, where $|X\rangle$ is an arbitrary linear-basis state and $|Y\rangle = |X + 90\rangle$ is the orthogonal-basis state. One photon goes to Alice and the other to Bob. (2) Alice chooses at random either to let her photon pass or to insert a half-wave plate (HWP), which will rotate her photon by $+45^\circ$. The latter choice changes the relative orientation between the two photons by $+45^\circ$. In the case shown, she chooses to rotate her photon. The new entangled state is $|\Phi'\rangle$. (3) Alice uses a polarizing

beam splitter (PBS) to measure her photon in the H/V basis. This optical element transmits horizontally polarized photons and reflects vertically polarized photons, and her unpolarized photon can collapse to either a horizontal or vertical polarization with equal probability. In this case, it collapses to a horizontal polarization. Alice records a bit value of 0. (5) Bob's photon was entangled with Alice's, so as a result of her measurement, his photon assumed the definite polarization state $|H - 45\rangle = |-45\rangle$. If Bob makes the same choice as Alice and inserts his HWP, he will rotate his photon's polarization by $+45^\circ$ and into a horizontal polarization. His photon will register in the H-detector,

and he will record a bit value of 0. If he makes the opposite choice and doesn't rotate his photon, the photon polarized at -45° has an equal probability of going to either detector (bit value either 0 or 1). As seen in Table I on the next page, whenever Bob and Alice make the same choice, they keep the bit because their bit values coincide. If they make opposite choices, they discard the bit since the values are not correlated. Alice and Bob can construct an identical sequence of random bits—a cryptographic key—simply by declaring their sequence of choices. The discussion can be public because the bit values are never revealed.

the key to each bit of her message, using modulo 2 addition ($0 + 0 = 0$, $0 + 1 = 1$, $1 + 0 = 1$, and $1 + 1 = 0$). She then sends the encrypted message to Bob, who decrypts it simply by repeating the operation, that is, by performing a sequential, bit-by-bit modulo 2 addition of the key to the message.

This type of encryption protocol, known as a one-time pad, is currently the only provably secure protocol. But the one-time pad is effective only if Alice and Bob never reuse the key, and more obviously, if the key remains secret. A potential eavesdropper, Eve, cannot be allowed to glean any part of

the bit stream that makes up the key. Therein lies a central problem of cryptography: How can secret keys be created and then securely distributed? The nonlocal correlations of entangled photons can play a role in this regard. (One can also exploit the properties of nonentangled photons in cryptographic schemes. See the article "Quantum Cryptography" on [page 68](#).)

In the entangled-state quantum cryptography scheme, Alice and Bob perform an experiment similar to the one described in the first section of the paper. They use a source S_3 that emits entangled photons in the general

state $|\Phi_+\rangle = 1/\sqrt{2} (|XX\rangle + |YY\rangle)$, where $|X\rangle$ is an arbitrary linear-basis state and $|Y\rangle$ is the orthogonal-basis state. One photon goes to Alice and the other to Bob. In this protocol, however, either banker can choose—at random and independent of each other—to use a half-wave plate (HWP) to rotate photon polarization by a set amount. The bankers then detect the photon in the H/V basis using a polarizing beam splitter, which transmits horizontally polarized photons and reflects vertically polarized photons (see Figure 3). Detection of a horizontally polarized photon is recorded as a 0; of a verti-

Table I. Constructing a Cryptographic Key with Entangled Photons

First Receiver (Alice)			Polarization to Second Receiver	Second Receiver (Bob)			Communication Results
Angle of Rotation (°)	Detector	Bit Value		Angle of Rotation (°)	Detector	Bit Value	
0	H	0	H	0	H	0	Keep bit
0	V	1	V	0	V	1	Keep bit
0	H	0	H	+45	H or V	0 or 1	Discard bit
0	V	1	V	+45	H or V	0 or 1	Discard bit
+45	H	0	−45°	0	H or V	0 or 1	Discard bit
+45	V	1	+45°	0	H or V	0 or 1	Discard bit
+45	H	0	−45°	+45	H	0	Keep bit
+45	V	1	+45°	+45	V	1	Keep bit

cally polarized photon, as a 1.

After a sufficient number of measurements (that number is dictated by the length of the key), Alice and Bob have a public discussion, during which they reveal whether they inserted the HWP before each measurement. At no time do they reveal the actual measurement results. Whenever Alice and Bob make the same choice (50 percent of the time), they know from the properties of entangled photons that they will have completely correlated results. By contrast, if one of them uses the HWP and the other doesn't, they will discard the results because their measurements would be completely uncorrelated (see Table I). Following this public discussion, each banker will be able to privately construct the same random string of 0s and 1s—an ideal key for cryptography.

What about the eavesdropper Eve? She is completely foiled in her attempts to know the secret key. Certainly, she cannot tap the photon line, as she might with conventional, classical communications. A single, indivisible quantum object—namely, a photon—is the conveyor of information in this cryptographic protocol. If Eve steals Bob's photon (a “denial-of-service” attack), the photon's information never becomes part of the key. Thus, although a wiretap

would reduce the rate of the transmission, it would not jeopardize the security of the key.

Eve can try to intercept the photon, measure it, and send another one to Bob. But any measurement Eve would make to determine the photon's polarization state would necessarily perturb the photon and collapse the entangled state. The photon she sends to Bob would therefore be classically correlated with Alice's photon. Consequently, Eve's intervention would necessarily induce additional errors into Bob's key.

This last point is significant. Unlike their theoretical counterparts, the encryption keys created by an actual quantum cryptography system initially have a small fraction of errors, because real equipment is always less than perfect. To make sure their key is secure, Alice and Bob ascribe all errors to Eve and then use this “bit error rate” to estimate the maximum amount of information available to the eavesdropper. They then use a privacy amplification scheme (discussed in the cryptography article on [page 68](#)) to reduce Eve's knowledge of the secret key to less than one bit.

But the bit error rate alone can lead to a false sense of security. If nonentangled photons with a definite polariza-

tion are sent to Bob, it is conceivable that some other degree of freedom may also be coupled to the polarization state. For example, if separate lasers are used to produce the two polarization states, the photons from each laser may have slightly different timing characteristics or frequency spectra. Such a difference would in principle allow an eavesdropper to distinguish between photons without disturbing the polarization state and, hence, without affecting the bit error rate.

When the photons are entangled, however, any leakage of information to other degrees of freedom can be shown to automatically manifest itself in the error rate detected by Alice and Bob. In other words, any degree of freedom with which the polarization might be coupled will cause noticeable effects on the polarization correlations. Therefore, using only the detected error rates, one can set an upper limit on the information available to an eavesdropper, even one who is not directly measuring the polarization of the photons, and then use privacy amplification to eliminate that information.

As a last resort, Eve may think of “cloning” Alice's photon. She could measure the clone while allowing the original to continue on to Bob, thus completely covering her tracks. But she

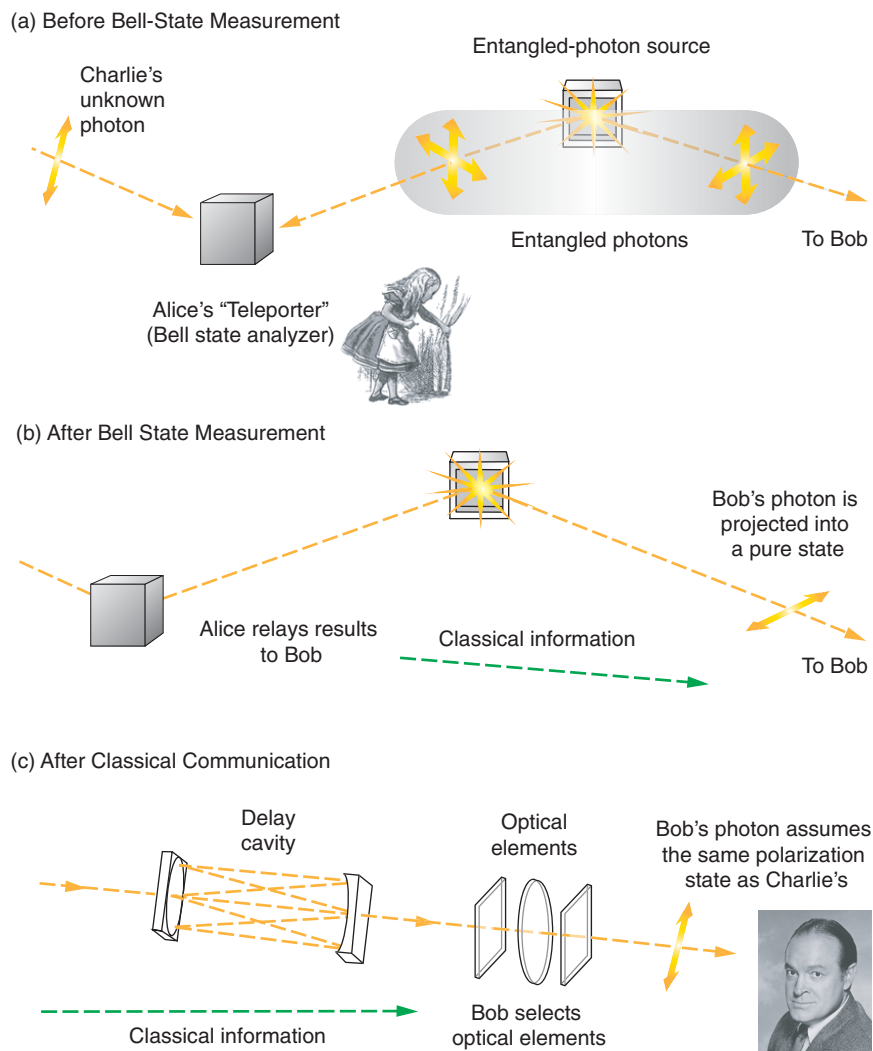


Figure 4. Quantum State Teleportation

(a) Alice's teleportation lab consists of an entangled photon source and a Bell state analyzer (the teleporter). One entangled photon goes to Bob and the other to the teleporter. Charlie sends a photon of unknown polarization state into the teleporter. (b) Alice performs a joint polarization measurement of the two photons in the teleporter and relays the result to Bob using two classical bits of information. The photon going to Bob is projected into a pure state as a result of Alice's measurement. (c) Upon receiving Alice's classical information, Bob performs a simple transformation on his photon, such as a rotation of the polarization vector. He duplicates the polarization state of Charlie's photon without knowing anything about its original state.

is again foiled by quantum mechanics. According to the no-cloning theorem, it is impossible to make a copy of a photon in an unknown state while simultaneously preserving the original. (See the box "The No-Cloning Theorem" on [page 79](#).) Eve is clearly out of business.

Teleportation. In 1993, Charles Bennett of IBM, Yorktown Heights, and his colleagues proposed a remarkable experiment with entangled particles, namely, the "teleportation" of a pure quantum state from one location to another.

Charlie wants to send his friend Bob a photon in an arbitrary, pure

quantum state $|\psi\rangle = \alpha|H\rangle + \beta|V\rangle$. He enlists the aid of Alice, who happens to run the Teleportation Laboratory shown in Figure 4. Inside the lab, a source S_3 is emitting a pair of entangled photons, one of which goes off to Bob. The other photon is input into Alice's "teleporter." Charlie is instructed to send his photon into the teleporter as well.

Alice then performs a special joint measurement of the polarization state of the two photons in the teleporter. She relays the result to Bob, who subsequently performs a simple transformation of the polarization state of his photon. As if by magic, the state of Bob's photon is transformed into

the state of Charlie's original photon.

Mathematically, this magic is described as follows. The three-photon initial state (that is, Charlie's photon plus the two entangled photons) can be represented as

$$|\psi_0\rangle = (\alpha|H\rangle + \beta|V\rangle)_C \times 1/\sqrt{2}(|HH\rangle + |VV\rangle)_{A,B}, \quad (1)$$

where the subscripts C, A, and B refer to Charlie's, Alice's, and Bob's photons, respectively. But $|\psi_0\rangle$ can also be represented as a superposition of states, each constructed in the following way: Charlie and Alice's photons are represented by one of the Bell states $|\Phi_{\pm}\rangle = 1/\sqrt{2}(|HH\rangle \pm |VV\rangle)$ and

$$|\psi_{\pm}\rangle = 1/\sqrt{2}(|HV\rangle \pm |VH\rangle),$$

and Bob's photon is represented as a photon in a pure state. Thus,

$$\begin{aligned} |\psi_0\rangle = & 1/2\{|\Phi_{-}\rangle_{C,A}(\alpha|H\rangle - \beta|V\rangle)_B \\ & + |\Phi_{+}\rangle_{C,A}(\alpha|H\rangle + \beta|V\rangle)_B \\ & + |\Psi_{-}\rangle_{C,A}(-\beta|H\rangle + \alpha|V\rangle)_B \\ & + |\Psi_{+}\rangle_{C,A}(\beta|H\rangle + \alpha|V\rangle)_B\} \quad (2) \end{aligned}$$

Technically speaking, this representation is possible because the Bell states are a basis for the two-photon Hilbert space and any state of two photons can be represented as a linear superposition of these states. It is important to point out that Alice's photon remains entangled with Bob's. Teleportation relies on Alice's ability to perform a joint polarization measurement that explicitly projects the two photons in the teleporter into one of the four Bell states. Once Alice completes her measurement, Bob's photon (which is totally correlated to Alice's) will assume the corresponding pure state. For example, if the Bell state measurement produces the result $|\Psi_{-}\rangle_{C,A}$, then Bob's photon would be projected into the pure state $|\psi\rangle = (-\beta|H\rangle + \alpha|V\rangle)_B$. By using a simple optical element, Bob can rotate the polarization state of his photon by 90° and transform it into the state $|\psi'\rangle = (\alpha|H\rangle + \beta|V\rangle)_B$, that is, the original input state. Provided Alice can specify which Bell state was measured (a specification that requires two bits of classical information), Bob can always choose an appropriate optical element to effect the proper rotation.

In a series of groundbreaking experiments conducted at the University of Innsbruck, Austria, Anton Zeilinger

and coworkers were the first to demonstrate quantum teleportation. The group is now able to determine two of the four Bell states unambiguously (the other two states give the same experimental signature²) and prove for those cases that the state of Charlie's photon could indeed be transferred to Bob's.

Several points should be made about quantum teleportation. First, during the entire procedure, neither Alice nor Bob has any idea what the values are for the parameters α and β that specify Charlie's photon. The initial, arbitrary pure state remains unknown. Second, teleportation is not cloning. The original state of Charlie's photon is necessarily destroyed by Alice's measurement, so the photon that Bob ends up with is still one of a kind.

Finally, hopeful sci-fi fans may be a little disappointed by this realization of teleportation. Unlike the TV show "Star Trek," in which Captain Kirk could be transported body and soul from the starship *Enterprise* to the surface of an alien planet,³ here only certain information about the photon is transferred to a photon in some faraway location. Because photons have numerous degrees of freedom in addition to their polarization, the original and the teleported photons are two different entities. And it goes without saying that an even simpler way for Charlie to send his quantum state to Bob would be to dispatch the original photon directly to him.

Nevertheless, teleportation remains an interesting application of quantum state entanglement. Furthermore, researchers have discussed how it might form the basis of a distributed network of quantum communication channels and how this basic informa-

tion protocol might be useful for quantum computing.

Quantum Microscopy and

Lithography. The general topic of quantum metrology involves capitalizing on the ultrastrong correlations of entangled systems to make measurements more precisely than would be possible with classical tools. The two main photon-based applications under investigation are quantum microscopy and quantum lithography.

At present, two-photon microscopy is widely used to produce high-resolution images, often of biological systems. However, the classical light sources (lasers) used for the imaging have random spreads in the temporal and spatial distributions of the photons, and the light intensity must be very high if two photons are to intersect within a small enough volume and cause a detectable excitation. The high intensity can damage the system under investigation. Because the temporal and spatial correlations may be much stronger between members of an entangled photon pair, one could conceivably get away with much weaker light sources, which would be much less damaging to the systems being observed. Moreover, entangled-photon sources may also enable obtaining enhanced spatial resolution.

Lithography, in which a pattern is optically imaged onto some photoresist material, is the primary method of manufacturing microscale or nanoscale electronic devices. An inherent limitation of this process is that details smaller than a wavelength of light cannot be written reliably. However, quantum state entanglement might circumvent this limitation. Under the right circumstances, the interference pattern formed by beams of entangled photon pairs can have half the classical fringe spacing.

Quantum lithography requires two beams of photons, which we will call A and B, but in this case, the type of entanglement is different

² Distinguishing between the four Bell states is still an unsolved technical problem. It requires a strong nonlinear interaction between two photons, which is extremely difficult to achieve in practice.

³ "Teleportation" (though it was not explicitly called that) was supposedly introduced in this TV show because the producer, Gene Roddenberry, wished to save the expense of simulating the landing of a starship on a planet.

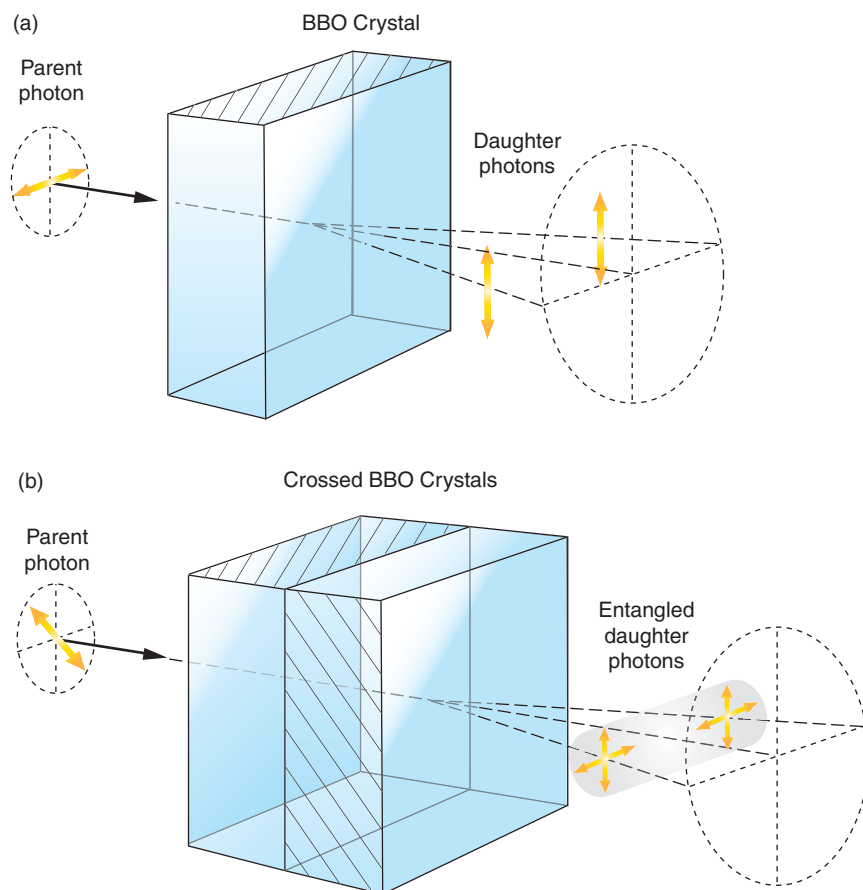


Figure 5. Entangled-Photon Source

(a) For a given orientation of the beta-barium borate (BBO) crystal, a horizontally polarized parent photon produces a pair of vertically polarized daughters. The daughters emerge on opposite sides of an imaginary cone. The cone's axis is parallel to the original direction taken by the parent photon. The two daughter photons are not in an entangled state. Reorienting the BBO crystal by 90° will produce a pair of horizontally polarized daughters if a vertically polarized pump beam is used. (b) Passing a photon polarized at $+45^\circ$ through two crossed BBO crystals can produce two photons in an entangled state. Because of the Heisenberg uncertainty principle, there is no way to tell in which crystal the parent photon "gave birth," and so a coherent superposition of two possible outcomes results: a pair of vertically polarized photons or a pair of horizontally polarized photons. The photons are in the maximally entangled state $|\Phi_+\rangle = 1/\sqrt{2}(|HH\rangle + |VV\rangle)$.

from the one discussed in the previous sections. What is needed is a coherent superposition consisting of the state in which two photons are in beam A while none are in B and the state in which no photon is in beam A while two photons are in B. Such number-entangled states can be made in the

laboratory, and the predictions about fringe spacings have been verified. However, other obstacles must be overcome in order to surpass current classical-lithography techniques. Researchers continue to explore the potential of this idea with the hope of achieving a viable commercial technology.

Creating and Measuring Entangled States

If quantum state entanglement is such a remarkable property because it allows one to perform secret communications, teleport states, or test the nonlocality of quantum mechanics, one naturally wonders how to make entangled states. Currently, scientists can create entangled states of particles in a controlled manner by using several technologies such as ion traps, cavity quantum electrodynamics (QED), and optical down-conversion. Here, we will concentrate on the optical realization.

Crystals of a certain chemical structure, such as beta-barium borate (BBO), have the property of optical nonlinearity, which means that the polarizability of these crystals depends on the square (or higher powers) of an applied electric field. The practical upshot of this property is that, when passing through such a crystal, a single-parent photon can split (or down-convert) into a pair of daughter photons. The probability that this event occurs is extremely small; on average, it happens to only one out of every 10 billion photons!

When down-conversion does occur, energy and momentum are conserved (as they must be for an isolated system). The daughter photons have lower frequencies (longer wavelengths) than the parent photon and emerge from the crystal on opposite sides of a cone that is centered about the direction traveled by the parent. For what is known as Type I phase matching, the daughters emerge from a specifically oriented BBO crystal with identical polarizations that are aligned perpendicular to the parent polarization—see Figure 5(a). Because each photon is in a definite state of polarization, the two photons are not in an entangled state but are classically correlated.

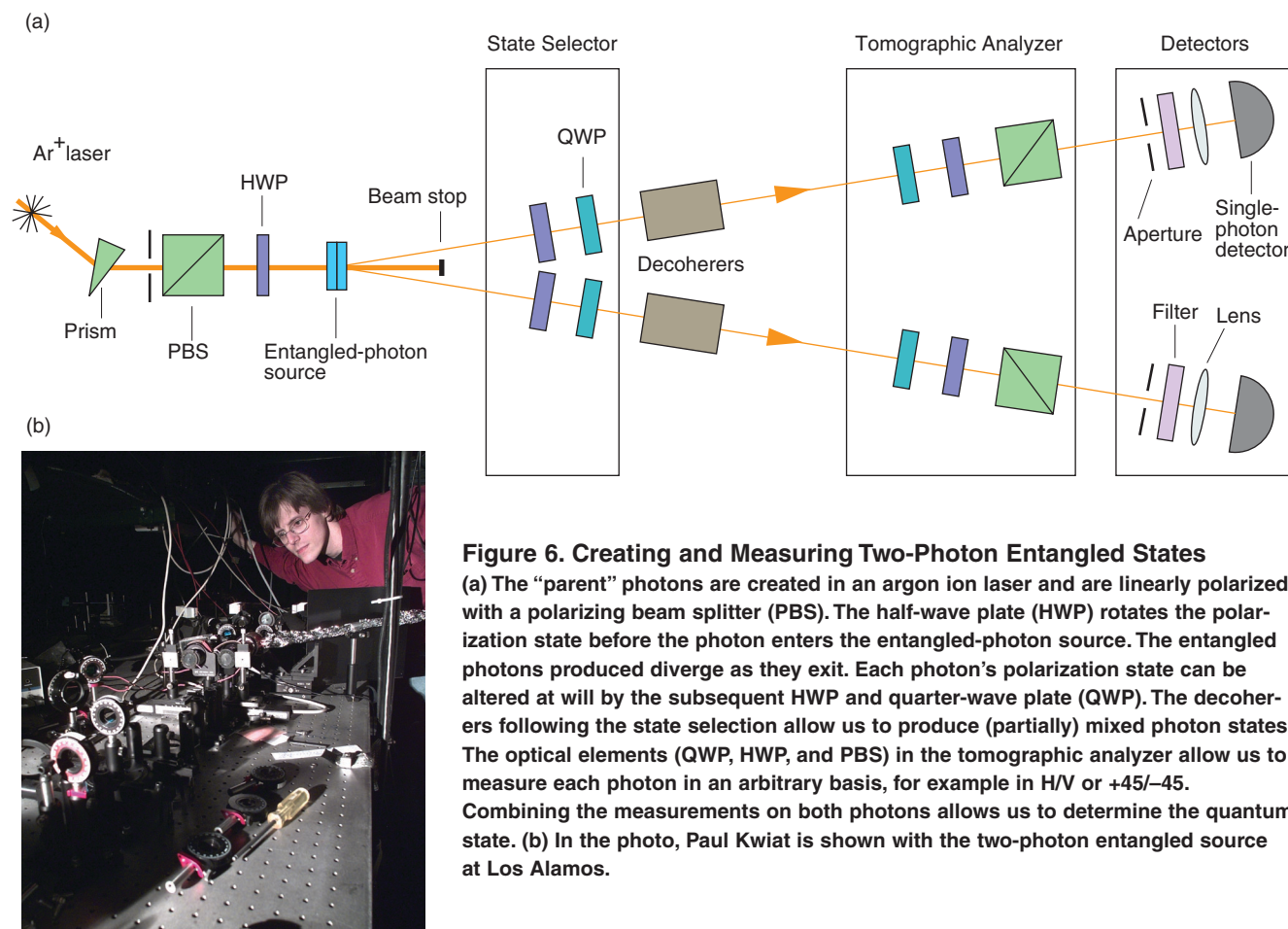


Figure 6. Creating and Measuring Two-Photon Entangled States

(a) The “parent” photons are created in an argon ion laser and are linearly polarized with a polarizing beam splitter (PBS). The half-wave plate (HWP) rotates the polarization state before the photon enters the entangled-photon source. The entangled photons produced diverge as they exit. Each photon’s polarization state can be altered at will by the subsequent HWP and quarter-wave plate (QWP). The decoherers following the state selection allow us to produce (partially) mixed photon states. The optical elements (QWP, HWP, and PBS) in the tomographic analyzer allow us to measure each photon in an arbitrary basis, for example in H/V or +45/−45. Combining the measurements on both photons allows us to determine the quantum state. (b) In the photo, Paul Kwiat is shown with the two-photon entangled source at Los Alamos.

(The crystal acts like the source S_1 described earlier.)

To create photons in the entangled state, one can use two crystals that are aligned with their axes of symmetry oriented at 90° to each other, as shown in Figure 5(b). With crossed crystals, two competing processes are possible: The parent photon can down-convert in the first crystal to yield two vertically polarized photons ($|VV\rangle$), or it can down-convert in the second to yield two horizontally polarized photons ($|HH\rangle$). It is impossible to distinguish which of these processes has occurred. Thus, the state of the daughter photons is a coherent quantum-mechanical superposition of the states that would arise from each crystal alone; the crossed crystals

produce photons in the state $|\Psi_{\text{out}}\rangle = 1/\sqrt{2}(|HH\rangle + |VV\rangle)$, which is maximally entangled.⁴

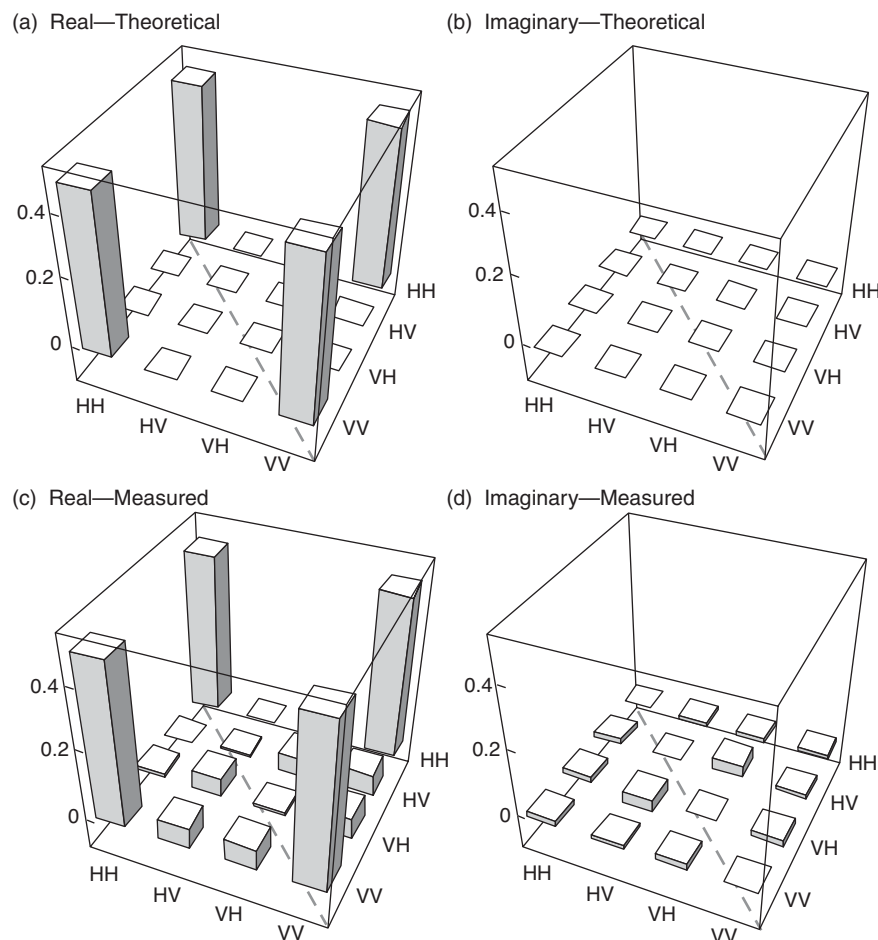
Figure 6 shows how this basic source can be adapted to produce any pure quantum state of two photons by placing rotatable half- and quarter-wave plates (which can be used to transform the polarization state of a single photon) before the crystal and in the paths of the two daughter photons. To create more general quantum

states—mixed states—a long birefringent crystal can be used to delay one polarization component with respect to the other. If the relative delay is longer than the coherence time of the photons, the horizontal and vertical components have been effectively decohered; that is, the phase relationship between the different states is destroyed. Researchers are still discovering how to combine sources and polarization-transforming elements to create all possible two-photon quantum states.

⁴ In an alternative approach known as “Type II phase matching,” only one crystal is needed to create the entangled state. The crystal has a different orientation, and each of the daughter photons emerges from the crystal on one of two possible exit cones. Entangled photons created by this approach were used in the first demonstration of quantum teleportation.

Characterizing Entanglement: The Map of Hilbert Space

As discussed in the box on [page 56](#), a mixed state of two photons (or in general, a mixed state of two qubits)

**Figure 7. Density Matrices**

Theoretical and experimental density matrices for the entangled state $|\Phi_+\rangle = 1/\sqrt{2} (|HH\rangle + |VV\rangle)$ are illustrated here. Both real and imaginary parts of the matrix are shown. The value of each matrix element is derived from the results of thousands of two-photon correlation experiments (simulated experiments for the theoretical matrix.) The experimental matrix indicates that our source can output a state close to a maximally entangled one. Written out “longhand,” the density matrix describing the state $|\Phi_+\rangle$ is

$$\begin{aligned} \rho &= |\Phi_+\rangle\langle\Phi_+| \\ &= 1/2(|HH\rangle\langle HH| + |VV\rangle\langle VV| \\ &\quad + |HH\rangle\langle VV| + |VV\rangle\langle HH|) . \end{aligned}$$

The first two terms, which lie on the diagonal of the matrix (dashed line), give the probability of the result (for example, 50% HH and 50% VV). The other two terms describe the quantum coherence between the states $|HH\rangle$ and $|VV\rangle$. For a classical mixed state (such as the source S_2 described in the text), these off-diagonal terms in the density matrix would equal zero. Notice that all coefficients in this density matrix are real, so that all terms in the imaginary part of the matrix should be zero.

is represented by a 4×4 density matrix, which is described by 15 independent parameters (15 real numbers). To determine the independent parameters, we make 15 coincidence measurements on the ensemble of photon pairs emitted from the source. Each measurement is similar to the one used in the simple experiment described at the start of this article. The measurement may be made with the tomographic analyzer shown in Figure 6. Using such a system, we were able to determine the density matrices of many types of states. An example is shown in Figure 7.

Whereas 15 numbers fully describe a two-photon mixed state, the density matrix for N photons needs $4^N - 1$ real numbers. Thus, the density matrix of a 4-photon state contains

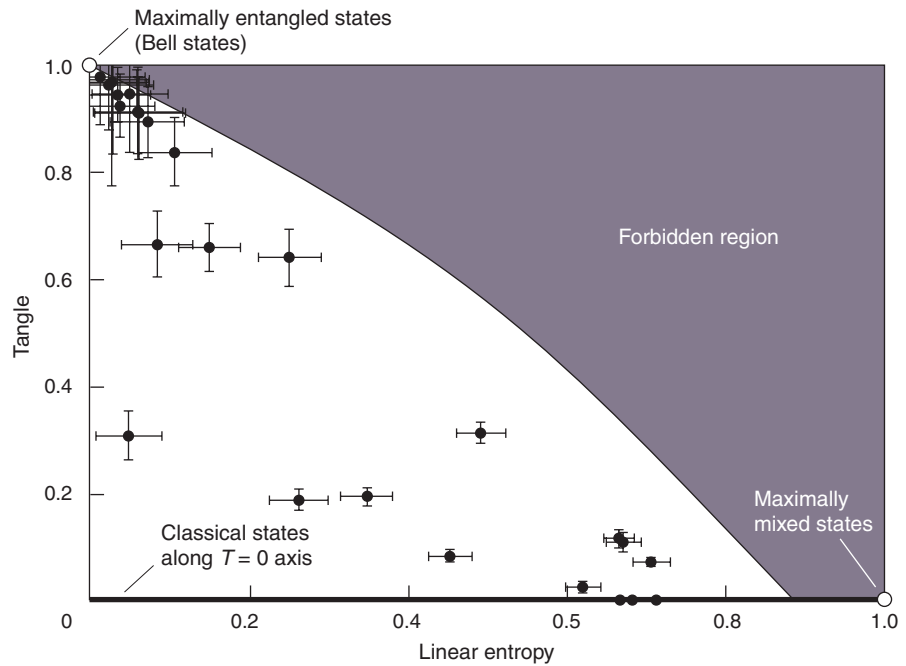
255 parameters and requires 255 separate measurements just to characterize the state. Note that, if each parameter is allowed to assume one of, say, 10 possible values, those 4 photons can be in any of 10^{255} distinct quantum states! This number of states is many orders of magnitude greater than the total number of particles in our universe. The mathematical space in which the quantum states rest (the Hilbert space) is unfathomably large, and in order to have any hope of navigating it, one needs to introduce a simpler representation for quantum states.

Two characteristics of central importance for quantum information processing are the extent of entanglement and the degree of purity of an arbitrary state. A quantity

called the von Neumann entropy has been introduced to characterize the degree of purity. (See the box “Characterizing Mixed States” on the [next page](#).) However, for the analysis of two-photon states, we found it easier to use a related quantity, known as the linear entropy. When the linear entropy equals zero, the state is pure. When it reaches its maximum value of 1, the state is completely random.

Measuring the entanglement of a mixed state is more complicated and, in general, is an unsolved research problem when more than two qubits are involved. Any mixed quantum state can be thought of as an incoherent combination of pure states: The system is in a number of possible pure states, each of which has some probability

Figure 8. The Map of Hilbert Space
The amount of entanglement (or the tangle) is plotted against the degree of purity (represented by the linear entropy) for a multitude of two-photon states created and measured at Los Alamos. Each state is represented by a black spot with error bars. The boundary line, which represents the class of states that have the maximum possible entanglement for a given value of the linear entropy, was first determined theoretically but then confirmed by a numerical simulation of two million random density matrices. Important states, such as those that are maximally entangled or completely mixed, are indicated. Efforts are under way to create states that lie along the boundary line.



between 0 and 1 associated with it (rather than the complex numbers defining the probability amplitudes that specify a particular superposition of pure states). A reasonable measure of the entanglement of such a mixed state is to take the average value of the entanglement (for example, as measured by the concurrence discussed in the box on this page) for all those pure states.

One must, however, use this procedure carefully because the decomposition of the mixed state into an incoherent sum of pure states is not unique. For this “average entanglement” to make any sense as a measure of entanglement of the mixed state, one must use the decomposition for which the average is a minimum. The square of this minimized quantity is called the “tangle.” It has a value of zero for entirely unentangled, separable states and of unity for completely entangled states.

Figure 8 shows how those two parameters—tangle and linear entropy—can be used to create a simplified map of Hilbert space for two-photon states. The crosses (with error bars) are the states we

Characterizing Mixed States

It is convenient to characterize the extent of entanglement and the degree of purity of a mixed state using two derived parameters: the tangle and the linear entropy. The linear entropy, which gives a measure of the purity of the state, derives from the von Neumann entropy. The latter is given by the formula $S = -\text{Tr}\{\rho \log_2(\rho)\}$, where ρ is the density matrix. Here $\text{Tr}\{M\}$ is the trace of a matrix (that is, the sum of terms on the diagonal) and \log_2 is a logarithm base 2, which can be defined for matrices via a power series. The von Neumann entropy is zero for a pure state. When the von Neumann entropy has its maximum value (equal to the number of qubits), the state is completely random, with no information or entanglement being present. The linear entropy, defined for two qubits as $S_L = 4/3(1 - \text{Tr}\{\rho^2\})$, is similar to the von Neumann entropy, but it is easier to calculate. Specifically, it equals 0 for a pure state and has a maximum value of 1 for completely random states.

Characterizing the degree of entanglement is more difficult. Mathematically speaking, if one decomposes the density matrix into an incoherent sum of pure states, that is, $\rho = \sum_i p_i |\psi_i\rangle \langle \psi_i|$, where $0 \leq p_i \leq 1$ and $\sum_i p_i = 1$, then the average entanglement is $\bar{E} = \sum_i p_i C(\psi_i)$, where $C(\psi_i)$ is the concurrence of the pure state $|\psi_i\rangle$ (defined in the box on page 56). It is very important to find the decomposition for which \bar{E} takes its minimum possible value; otherwise, one can infer a nonzero entanglement for states such as the completely mixed state, which is certainly not entangled! Fortunately, the way to do that decomposition has been worked out for two qubits. Characterizing the degree of entanglement for three or more qubits remains an unsolved research problem.

have created and measured experimentally. Most display a high degree of entanglement. States created by other technologies can be plotted on such a diagram as well.

Conclusions

Entangled states arise naturally whenever two or more quantum systems interact. In fact, one of the prevalent theories of nature is that the universe is really one big, vastly complicated entangled state, described by the “wave function of the universe.” Despite their seeming ubiquity, however, entangled states are not generally observed in the world at large. Only relatively recently have scientists developed the means to controllably produce, manipulate, and detect this most bizarre quantum phenomenon. Initially, the fascination was limited to experimental studies of the foundations of quantum mechanics, especially the notion of nonlocal “spookylike” influences (to quote Einstein). However, even more recently, has come the realization that entanglement could lead to enhanced—sometimes vastly enhanced—capabilities in the realm of information processing.

This paper has discussed how entangled states could be a key resource in applications as diverse as cryptography, lithography, and metrology because they enable feats beyond those possible with classical physics. In addition, the quest to create a quantum computer has pushed entangled systems to the forefront of quantum research. Part of the power of a quantum computer is that it creates entangled states of N qubits so that information can be stored and processed in the 2^N -dimensional qubit space. Quantum algorithms have been developed that would manipulate the complex entangled state and make use of the nonclassical correlations to solve problems more efficiently than

could be done classically. Scientists who work on developing quantum computers are envisioning systems of thousands of entangled qubits.

We don’t know whether we will be able to create or maintain such a complex entangled state. At this point, we won’t even claim to know whether we will fully understand that state if it is created. More research is needed before those questions can be answered. All that we can say now is that the once-hidden domain of quantum entanglement has broken into our classical world. ■

Further Reading

- Bouwmeester, D., A. K. Ekert, and A. Zeilinger, eds. 2000. *The Physics of Quantum Information*. Berlin: Springer-Verlag.
- Haroche, S. 1998. Entanglement, Decoherence and the Quantum/Classical Boundary. *Phys. Today* **51** (7): 36.
- James, D. F. V., P. G. Kwiat, W. J. Munro, and A. G. White. 2001. Measurement of Qubits. *Phys. Rev. A* **64**: 052312.
- Mandel, L. and E. Wolf. 1995. *Optical Coherence and Quantum Optics*. Cambridge: Cambridge University Press.
- Naik, D. S., C. G. Peterson, A. G. White, A. J. Berglund, and P. G. Kwiat. Entangled State Quantum Cryptography: Eavesdropping on the Ekert Protocol. 2000. *Phys. Rev. Lett.* **84**: 4733.
- Nielson, M. A., and I. L. Chuang. 2000. *Quantum Computation and Quantum Information*. Cambridge: Cambridge University Press.
- Schrödinger, E. 1935. Discussion of Probability Relations Separated Systems. *Proc. Cambridge Philos. Soc.* **31**: 555.
- White, A. G., D. F. V. James, P. H. Eberhard, and P. G. Kwiat. 1999. Nonmaximally Entangled States: Production, Characterization, and Utilization. *Phys. Rev. Lett.* **83** (16): 3103.
- Zeilinger, A. 2000. Quantum Teleportation. *Sci. Am.* **282** (4): 50.

Daniel F. V. James was born in rainy Manchester, in northwest England, within sight of the Old Trafford (home of Manchester United, the most successful sports team in recorded human history). He was educated at New College, University of Oxford, England, and at the Institute of Optics, University of Rochester, New York, where he earned his Ph.D. in optics in 1992. He came to Los Alamos National Laboratory as a postdoctoral fellow in 1994. Three years later, he became a staff member in the Atomic and Optical Theory Group at Los Alamos. Daniel specializes in theoretical optical physics, including coherence theory, diffraction, scattering, statistical optics, and quantum technologies.




Paul G. Kwiat comes from Ohio (not a place noted for good sports teams). He was educated at the Massachusetts Institute of Technology and the University of California at Berkeley,



where he received his Ph.D. in physics. In 1995, after having completed a Lise Meitner postdoctoral fellowship at the University of Innsbruck, in Austria, Paul came to Los Alamos National

Laboratory as an Oppenheimer Fellow and became a staff member in 1998. In January 2001, he became the John Bardeen Professor of Physics and Electrical Engineering at the University of Illinois at Urbana-Champaign. He specializes in experimental quantum optics, with an eye toward shedding light on quantum information protocols.



The battle between cryptographers, who encrypt messages, and cryptanalysts, who break those codes, has raged for centuries. As quantum computing promises to help cryptanalysts break many of the encryption methods used today, quantum cryptography promises to keep our secrets safe forever.



A New Face for Cryptography

Jane E. Nordholt and Richard J. Hughes

Cryptography, the mathematical science of secret communications, has had a long and distinguished history dating back to the time of the ancient Greeks. It is a subject noted for the never-ending struggle for one-upmanship between code makers and code breakers, a struggle in which the future of nations has literally been at stake. The code breakers' need to read another party's secret communications has been a tremendous force driving the development of new information-processing technologies. The code makers have responded by using those new technologies to develop more complex methods for ensuring the security of communications.

The latest round in this struggle seems set to be played out in the world's physics laboratories, with the combatants drawing upon fundamental principles of quantum physics, principles that were only of academic interest until about 15 years ago. The code breakers believe that a large-scale quantum computer—a device that uses the nonclassical aspects of quantum systems to manipulate information—could defeat the most widespread cryptosystems in use today. They are pushing the physics community to develop such a computer,

which necessarily involves controlling atoms and photons in ways that were barely dreamed of—until recently. Meanwhile, the code makers are ready for battle and are already exploiting quantum mechanics in a new code-making technology—quantum key distribution (QKD)—that could counter the quantum computing threat.

Classical Cryptography

The main goal of cryptography is to allow two parties (conventionally referred to as “Alice” and “Bob”) to communicate while simultaneously preventing a third party (“Eve”) from understanding those communications. Alice and Bob's messages should remain secret even when Eve is able to passively monitor the exchanges. (A more intrusive Eve might want to prevent Alice and Bob from communicating at all, but such a denial-of-service attack is a different type of communication problem that we will not consider here.) Cryptography provides Alice with the means to render her messages to Bob in a form that is indistinguishable from random noise but that, nevertheless, allows Bob to recover the original message.

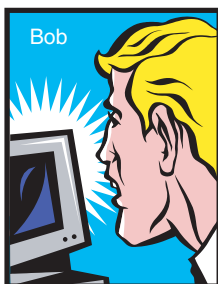
Figure 1. A Symmetric-Key Cryptography System: The One-Time Pad

(a) Alice, the sender, first generates a string of binary bits (the key) that is as long as her binary message. Then she applies the XOR operation—bit by bit—to the key and her message, and sends the encrypted string to Bob over an open communications channel. (b) Bob, the receiver, uses the same key as Alice to decrypt the message by the same XOR operation, applied bit by bit. His decrypted message is identical to the original message sent by Alice. Because the value of each key bit is random, the message cannot be recovered without the key. As long as Alice and Bob use the key only once to encrypt and decrypt one message, this one-time pad system is absolutely secure, but distributing the secret keys remains a problem. (c)–(e) This series of photographs shows an aerial view of the St. Louis International Airport before encryption, as encrypted by Alice, and as decrypted by Bob. Whereas Alice's encrypted photo is indistinguishable from random noise, Bob is able to reproduce the original faithfully.

(a) Encryption, One-Time Pad



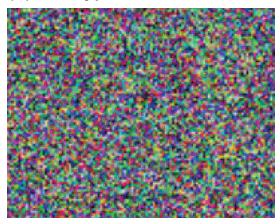
(b) Decryption



(c) Original



(d) Encrypted



(e) Recovered Original



XOR operation, \oplus :

$$0 \oplus 0 = 0 : 0 \oplus 1 = 1 : 1 \oplus 0 = 1 : 1 \oplus 1 = 0$$

Alice's message	1001	0000	0110	1001
\oplus				
Key	1000	0100	0101	0001
Encrypted message	0001	0100	0011	1000

Classical communication channel

Encrypted message	0001	0100	0011	1000
\oplus				
Key	1000	0100	0101	0001
Original message	1001	0000	0110	1001

This process of encryption (by Alice) and decryption (by Bob) can be accomplished if the two parties share a string of randomly generated binary bits known as a cryptographic key. In a system called the “one-time pad,” Alice and Bob must have identical copies of the key. (How they get the key will be discussed later). As seen in Figure 1, Alice adds the key to her message, bit by bit, using the binary exclusive OR- (XOR-, \oplus) operation, which is equivalent to addition modulo 2. Mathematically, the XOR operation is defined as

$$\begin{aligned} 0 \oplus 0 &= 0, \\ 0 \oplus 1 &= 1, \\ 1 \oplus 0 &= 1, \text{ and} \\ 1 \oplus 1 &= 0. \end{aligned} \quad (1)$$

Alice's encrypted communication at this point is indistinguishable from

random noise. Alice sends this message to Bob, who takes his copy of the key and subtracts it from the message, again using an XOR-operation. The original script is recovered. Provided a key is used to encipher only one message, the one-time pad encryption process is provably secure. In fact, it is the only completely secure cryptographic system.

The one-time pad is an example of a symmetric-key system (symmetric because Alice and Bob have the same key), and it requires a key that is as long as the message. In another type of symmetric key system, Alice and Bob use a short key to seed a high-quality random number generator of which they have identical copies. They then need to share fewer initial key bits in order to encrypt and decrypt large messages. In the Data Encryption Standard (DES)—a sym-

metric-key algorithm that was adopted as a United States government standard in 1977—the key length is 56 bits.

The security of all symmetric-key cryptographic systems rests entirely on the secrecy of the shared key because the structure of the cryptographic algorithm used by Alice and Bob is public knowledge. Certainly, the eavesdropper Eve understands and can implement the decryption algorithm. Should Eve obtain the key, she could immediately read Alice and Bob's messages. Without the key, Eve must attempt a mathematical attack on the encrypted message (or parts thereof) in order to crack it. In a properly designed symmetric-key cryptosystem, no attack should be more efficient than an exhaustive search over all possible keys.

Consider, for example, the 56-bit

DES key. Because there is a choice of either 0 or 1 for every bit in a binary key, there are 2^{56} (or nearly 10^{18}) possible DES keys. A desktop computer testing a million keys a second would require more than two thousand years to search the entire key space. But the phenomenal increase in computational speed and capability has made the 56-bit key vulnerable. Today's supercomputers can search all possible keys in a matter of hours.

The simple solution is to use longer keys. Adding a bit to the key length doubles the search time, whereas doubling the key length makes the search problem exponentially harder. In the forthcoming Advanced Encryption Standard (AES), the key length will be up to 256 bits, in which case a search of the entire key space would be so computationally demanding that it would not be feasible on any computer system within the useful lifetime of the information.

The Key Distribution Problem.

A DES-type cryptographic system reduces the act of communicating a long secret (the message) to that of creating and sending a short secret (the key). But the central issue within any system is that any information about the key must remain out of the hands of unwanted parties. This latter requirement creates what is known as the key-distribution problem.

Traditionally, cryptographic keys were distributed by trusted couriers immortalized in spy movies as strangers in trench coats handcuffed to locked briefcases. But the infrastructure required to manage the key material makes this type of distribution impractical in our computer-driven, global community. Picture the logistics nightmare if a courier had to deliver a cryptographic key every time Alice wanted to use her credit card over the Internet—and imagine the added cost! In some cases, courier key

distribution is even impossible, such as when Bob is not a person but a satellite in Earth's orbit. Furthermore, the existence of the key material before delivery by courier introduces an insider threat, in that the key material could be copied and delivered surreptitiously to Eve.

About 30 years ago, researchers at Britain's Government Communications Headquarters (GCHQ), and later (independently) in the United States, found a new, more convenient way to securely distribute cryptographic keys. The system is known generically as public-key cryptography. One public-key protocol begins when Bob generates two very large prime numbers, p and q , which are multiplied to form the especially large number N . He then selects an integer g , and uses the numbers p , q , and g to generate a fourth number, d . The two numbers (N, g) constitute Bob's public key, which he makes widely available. The number d constitutes Bob's private key, which he keeps secret. (The protocol is discussed in greater detail in the box "Public-Key Cryptography: RSA" on the [next page](#).)

When Alice wants to send an encrypted message to Bob, she grabs a copy of his public key and uses it in an algorithm that mathematically scrambles her communication. The algorithm, however, is a clever one-way operation: Bob's public key (N, g) cannot be used to unscramble Alice's encrypted message. Instead, one needs the secret number d from Bob's private key to decrypt. Given only N , it is extremely difficult to find the prime factors p and q that are needed to generate d ; hence, the system is considered secure.

Because the public-key cryptography system is asymmetric—only Bob needs to have a secret key—it has become the enabling technology for electronic commerce. Alice can grab the public key from the Bob.com website and safely encrypt and send

her credit card number. In addition, public-key encryption also provides a means for Alice to authenticate her transaction.

But public-key cryptography has its downside. Because of the computational difficulty in calculating asymmetric keys, Alice and Bob use it only to produce and distribute a symmetric key that they then use for the bulk of their discussions. More disturbing is the lack of proof that the methodology is secure. A clever person could come up with a new factoring algorithm that allows finding the secret number d , thus making public-key cryptography obsolete.

In 1994, Peter Shor of AT&T did invent such an algorithm. If implemented, that algorithm would undermine the public-key cryptography in use today. Fortunately, Shor's algorithm must be run on a quantum computer, which is currently unavailable and will probably remain so for many years.

Public-key cryptography clearly has a place where security need not be guaranteed to last for years. Because it is not provably secure, however, and because a quantum computer may render it useless in the future, a better system is needed for highly valuable data such as government or trade secrets. That better system is quantum cryptography.

Quantum Cryptography

Quantum cryptography is a type of symmetric-key distribution that allows Alice and Bob to create and share a secret key, while Eve is prevented from obtaining any more than a tiny fraction of one bit of information about the final key's binary sequence. The secret key can actually be used in any symmetric encryption method desired. Because quantum cryptography is used to send these key bits, it is more correctly called

Public-Key Cryptography: RSA

Public-key cryptography is an asymmetric key-distribution system, wherein Bob generates two keys: a public key, which he makes available to anyone, and a private key, which he keeps secret. Alice uses the public key to encrypt her message, which she then sends to Bob, who uses his private key to decrypt that message. Perhaps the most widely used public-key cryptography algorithm is RSA, which was invented in 1978 by Ron Rivest, Adi Shamir, and Leonard Adleman and was named for its inventors. The RSA algorithm uses two keys that are constructed as follows:

- Bob generates two prime numbers, p and q , which are typically very large (several hundred bits in length).
- He calculates the product, $N = pq$, known as the modulus.
- He calculates Euler's quotient function $\Phi(N)$, which is simply the number of integers less than N that are coprime* to N . If p is a prime number, every number less than p is coprime to it, so $\Phi(p) = p - 1$. Since the modulus $N = pq$ is the product of prime numbers, $\Phi(N) = (p - 1)(q - 1)$. Let $\Phi(N)$ be designated by η .
- Bob chooses an integer g such that $g < N$, and g has no factors in common with η .
- Bob calculates $d = g^{\Phi(\eta)-1} \bmod \eta$, where $\bmod \eta$ is the modulus operation.[†]

Bob's public key is (N, g) . His private key is the number d .

* Two integers are coprime if they share no common divisors except 1.

[†] For an introduction to modular arithmetic, see the article "From Factoring to Phase Estimation" on [page 38](#).

When Alice wants to send a message to Bob, she first represents her message as a series of numbers. To encrypt, she grabs Bob's public key (N, g) and uses it in the following mathematical transformation:

$$c = m^g \bmod N, \quad (1)$$

where m is a number representing a piece of her message. She sends the new number c off to Bob, who uses his private key (N, d) to perform the operation

$$m = c^d \bmod N, \quad (2)$$

thereby recovering Alice's number.

Public-key cryptography is based on a theorem by Euler, which states that $x^{\Phi(y)} = 1 \bmod y$, for any integer x that is coprime to the number y . The number d was chosen such that $d = g^{\Phi(\eta)-1} \bmod \eta$, or $dg = g^{\Phi(\eta)} \bmod \eta$, which by Euler's theorem becomes $dg = 1 \bmod \eta$. Subtracting 1 will result in $dg - 1 = 0 \bmod \eta$.

The last statement indicates that the number $dg - 1$ is evenly divisible by η , so that $dg - 1 = k\eta$, where k is an integer. In decrypting the message, Bob has

$$\begin{aligned} c^d \bmod N &= (m^g)^d \bmod N, \\ &= m (m^{dg-1} \bmod N), \text{ and} \\ &= m (m^{k\eta} \bmod N). \end{aligned} \quad (3)$$

But $\eta = \Phi(N)$. By Euler's theorem, $m^{\Phi(N)} = 1 \bmod N$. Thus,

$$\begin{aligned} c^d \bmod N &= m (1)^k \bmod N, \text{ and} \\ &= m \bmod N. \end{aligned} \quad (4)$$

In other words, $c^d \bmod N = m$, so that the decryption algorithm recovers Alice's message.

quantum key distribution (QKD). Adding to the security of a QKD system is the fact that any attempt to steal or copy a key can be detected, thus revealing information about the security environment.

The quantum part of quantum cryptography comes from the transmission and reception of single photons. In addition to keeping an eavesdropper at bay (primarily

because a photon cannot be split or copied reliably), quantum cryptographic systems exhibit strange quantum mechanical behaviors that are not normally observed in the classical world of everyday experience. The best example of such behavior occurs in our fiber-based quantum cryptographic system, in which we use the interference of single photons with themselves

to transmit information.

Before describing how a photon interfering with itself helps us encrypt messages, we will present an overview of the steps involved in executing a secure exchange of messages and then describe a simple protocol. Protocols are the rules used for the quantum mechanical and conventional transmissions at the heart of QKD.

A QKD Session. To perform QKD, Alice and Bob communicate in two different ways. The first is over a quantum channel, which allows Alice to reliably send single photons to Bob. While Eve may attempt to breach the quantum channel, her tampering can be detected. The second means of communication is an ordinary, public channel assumed to be monitored by Eve. Alice and Bob use this open channel to construct their secret key, implement any of several error-correction techniques, and coordinate a “privacy amplification” scheme that effectively prevents Eve from gaining any knowledge about the final key. In all, six steps are implemented in a QKD session. These are summarized in the box to the right.

As a first step, Alice and Bob authenticate their communications; that is, they verify each other’s identity. If this step is ignored, Eve can perform a “man-in-the-middle” attack and convince Alice that she is Bob, and Bob that she is Alice, in which case no form of key distribution or encryption can prevent Eve from reading all of Alice and Bob’s communications.

After authentication, Alice and Bob begin their QKD session. First, each generates a random bit stream. Alice then uses a QKD protocol, such as BB84 (discussed in the next section), that specifies how she is to encode each bit as the quantum state of a single particle. For example, she may use the specific polarization state of a single photon to encode for either a 0 or a 1. Then, Alice would send a stream of polarized photons to Bob, who follows the protocol in determining how to measure the polarization and hence deduce a bit sequence. Because of the way the protocol works, Alice and Bob can have a public conversation and select an overlapping subset of bits without revealing to each other the value of those bits.

Six Steps to a QKD Session

Authenticate. Over an open communication line, Alice confirms she is talking to Bob, and Bob confirms he is talking to Alice.

Use a quantum protocol. The protocol dictates how Alice is to encode her random bit stream as a quantum state of a single photon. Bob measures photons according to the protocol.

Construct the sifted key. Alice and Bob use an open line to discover which photons were sent and measured in the same basis. The bit values associated with that subset of photons form the sifted key.

Construct the reconciled key. Over the open line, Alice and Bob find and remove errors from the sifted key to make the reconciled key.

Construct the secret key. Alice and Bob use privacy amplification to construct a secret key from the reconciled key. An eavesdropper has essentially no information about the bits in the secret key.

Save some bits. A few secret bits are retained to enable authenticating future QKD sessions.

For example, if Alice’s random sequence is 0111 1010 1001 and as a result of his measurements Bob obtains the sequence 1001 1100 0100, then the protocol provides a means for Alice and Bob to know—without specifically telling each other—that the fourth, fifth, eighth, and eleventh bits form a common subsequence of 1100. This subsequence is called the “sifted” key.

In the real world, hardware is noisy, and transmission media are lossy, so the sifted key will contain some errors. Alice and Bob continue their public conversation and create a “reconciled” key, in which those errors are removed. During this process, some information about the sifted key becomes available to any potential listener (Eve). But Alice and Bob can calculate the maximum information Eve could have about their reconciled key, and using privacy amplification, reduce Eve’s information to substantially less than one bit. The result is a secret key known only to Alice and Bob. The one remaining step before closing the

session is to save a few key bits and thereby have a means to authenticate the next QKD session.

The BB84 Protocol. In 1984, Charles Bennett and Gilles Brassard published a paper describing how orthogonal and nonorthogonal quantum states could be used to construct a cryptographic key. Known today as BB84, the protocol is at the heart of our experimentally realized QKD systems. In the free-space version, Alice encodes random bit values in the polarization states of photons and then sends the single photons to Bob over the quantum channel. Bob’s measurement of the photon’s polarization and subsequent communication with Alice over a public channel allow the two parties to construct a sifted key.

A stylized version of the BB84 protocol is shown in Figure 2. (The box “Photons, Polarizers, and Projections” on [page 76](#) also provides some background material for this section.) Alice generates a random sequence of bits and then chooses—also at random—between one of two

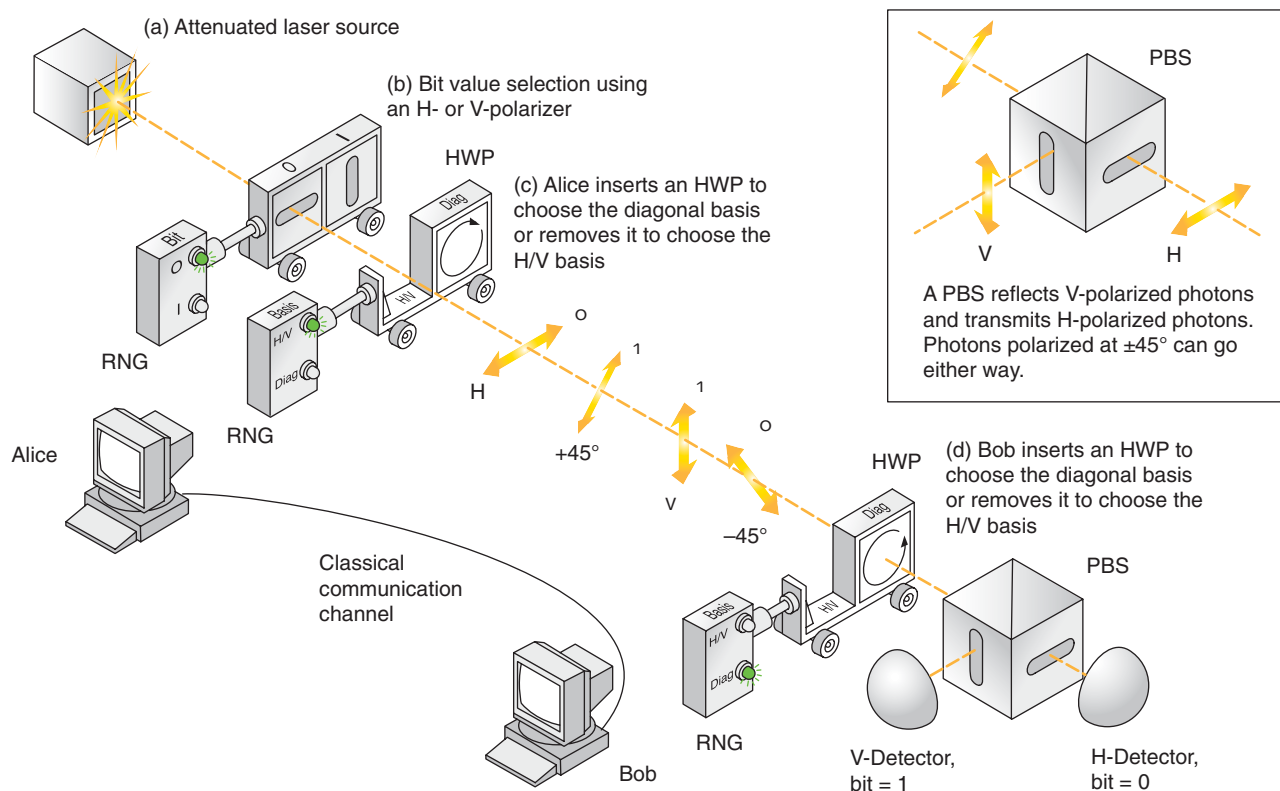


Figure 2. The BB84 Protocol

The BB84 protocol works because Alice randomly chooses to encode the photons in two, nonorthogonal bases. (a) An attenuated laser produces close to single photons. (b) Alice uses a random number generator (RNG) to select a bit value: 0s are encoded as horizontally polarized photons and 1s as vertically polarized photons (c) A second RNG selects the basis. To choose the H/V basis, Alice does nothing, (the photons are already either $|H\rangle$ or $|V\rangle$). To choose the diagonal ($-45^\circ/+45^\circ$) basis, she inserts a half-wave plate (HWP) that rotates the polarization by -45° , so that $|H\rangle$ goes to $|+45^\circ\rangle$ and $|V\rangle$ to $| -45^\circ\rangle$. (d) Bob uses an RNG to select his measurement basis, choos-

ing either to do nothing (H/V) or to rotate the photon by $\pm 45^\circ$ ($-45^\circ/+45^\circ$). He detects photons using an H/V oriented polarizing beam splitter (PBS), which transmits horizontally polarized photons but reflects vertically polarized ones (see inset). Photons polarized at $\pm 45^\circ$ have an equal probability to go to either detector. Table I shows that, when Alice and Bob choose the same basis, they know that their bit values coincide. When they choose different bases, their bit values are randomly correlated. At the end of the session, Bob and Alice openly compare their bases for each measurement. They keep only those bits that were sent and measured in the same basis.

Table I. Details of the BB84 Protocol

Sender (Alice)			Receiver (Bob)				Joint Action	
Alice's Basis	Bit	Polarization	Bob's Basis	Resulting Polarization	Probability (%)		Bit	
					H-Det.	V-Det.		
H/V	0	H	H/V	H	100	0	0	Keep bit
H/V	1	V	H/V	V	0	100	1	Keep bit
H/V	0	H	Diag.	$+45^\circ$	50	50	0 or 1	Discard bit
H/V	1	V	Diag.	-45°	50	50	0 or 1	Discard bit
Diag.	0	-45°	H/V	-45°	50	50	0 or 1	Discard bit
Diag.	1	$+45^\circ$	H/V	$+45^\circ$	50	50	0 or 1	Discard bit
Diag.	0	-45°	Diag.	H	100	0	0	Keep bit
Diag.	1	$+45^\circ$	Diag.	V	0	100	1	Keep bit

polarization bases, either the horizontal/vertical (H/V) basis, or the diagonal ($-45^\circ/+45^\circ$) basis. If she chooses the H/V basis, the bit values of 0 are encoded as horizontally polarized photons, and bit values of 1 are encoded as vertically polarized photons, that is, $0 = |H\rangle$ and $1 = |V\rangle$. Similarly, if she chooses the diagonal basis, 0 and 1 bit values are encoded as $0 = |-45\rangle$ and $1 = |+45\rangle$. She sends the stream of polarized photons off to Bob.

At his end, Bob chooses at random to measure polarizations in either the H/V or diagonal basis. As shown in Figure 2, he uses a special dual-detector system. If he chooses the H/V basis, then photons in the state $|H\rangle$ go through to his H-detector, while those in the state $|V\rangle$ are reflected to the V-detector. Photons in the $|-45\rangle$ or $|+45\rangle$ state go randomly to either detector. If Bob measures in the diagonal basis, then his setup directs $|-45\rangle$ photons to the H-detector, $|+45\rangle$ photons to the V-detector, and $|H\rangle$ or $|V\rangle$ photons to either detector with equal probability.

Table I shows how the results differ depending on which polarization states were sent and how they were detected. When Alice and Bob used the same basis, a photon hit on Bob's H-detector means that Alice had a bit value of 0; a hit on his V-detector, that she had a bit value of 1. If the bases differ, there is no such correspondence. Bob and Alice therefore use the public channel and simply compare the sequence of bases. They keep the corresponding bits when the bases agree and disregard the bits when they don't agree. In this way, they can build a sifted-key sequence over a public channel without ever revealing the value of the individual key bits.

Because Alice and Bob have a 50 percent chance of choosing the same basis, in an ideal implementation of BB84, half of the photons are used to create the sifted key. In practice, the efficiency is much less

because the real world unavoidably introduces errors into the sifted-key sequence—polarizers are not perfect, photons do not always reach Bob, and detectors do not always fire when hit with a photon and sometimes fire on their own. Alice and Bob must check and correct their sequence for errors.

Error Correction. One example of a simple error-correction scheme is illustrated in Figure 3. Alice tells Bob the parity of each of her bytes, that is, whether the sum of each 8 bits of the sifted key is even or odd. Bob then checks the parity of his bytes. They keep those bytes that have the same parity and initiate a 20-questions-type deductive process to find the problem bit when the parity differs.¹ Because parity checks can only find an odd number of errors in a bit sequence, in practice, sifted bits are shuffled and then checked for errors several times. All errors must be eliminated to a high degree of certainty. If Alice and Bob's keys differ by even a single bit, the keys will be unusable.

Alice and Bob make their byte comparisons over the open channel, so Eve now has—at a minimum—information about the parity of each retained byte. To eliminate even this limited knowledge on Eve's part, Alice and Bob can agree to drop the last bit of each byte. In addition, they have to sacrifice some key bits to find the errors in their sequences. The reconciled key is therefore shorter than the sifted key. While undertaking the error correction process, however, Alice and Bob obtain an estimate of the bit error rate (BER), which is the number of errors they had in their sifted sequences. Alice and Bob use the BER and knowledge of the quantum mechanical and physical principles of the QKD technique to put a rigorous upper bound on the possible information that Eve may have about their bit sequences.

Privacy Amplification. In this step, Alice and Bob do an XOR operation on sequences of bits from the reconciled key to produce fewer, but brand new, bits. The amount of compression required depends on their estimate of Eve's acquired knowledge.

For example, suppose Alice and Bob share a reconciled sequence consisting of six bits, a, b, c, d, e, and f, and they suspect that Eve knows three of the six bits. Alice and Bob make two new bits out of the original six by doing the following operation:

$$\begin{aligned} a \oplus b \oplus c \oplus d &= \text{Bit 1} \quad , \text{ and} \\ c \oplus d \oplus e \oplus f &= \text{Bit 2} \quad . \end{aligned} \quad (2)$$

Although Eve may have known three bits of the reconciled key sequence, she knows nothing about the new bits generated by privacy application. Alice and Bob can apply this procedure to reduce Eve's knowledge to less than one bit in a key that is several hundred bits long and thereby produce a completely secure key. In general, if the original sequence is n -bits long, privacy amplification will compress it to $R(n)$ bits, where

$$R(n) = -n \log_2[\zeta^2 + (1 - \zeta)^2] \quad (3)$$

and ζ is the BER.

Foiling Eve. We are now in a better position to discuss how the complete QKD session prevents Eve from gaining information about the secret key. First, Eve cannot get any information about the key over the open channel; although the BB84 protocol allows her to know which bits Alice and Bob had in common, she knows nothing about the values

¹ Bits that get transmitted correctly are valuable. Although Alice and Bob could drop all eight bits of a problem byte, it is usually worthwhile to winnow through the byte and retain as many bits as possible.

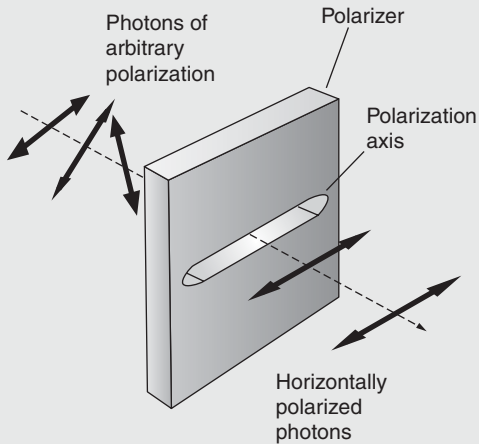


Figure A. Polarizing Filter
The filter projects photons into polarization states parallel to its polarization axis.

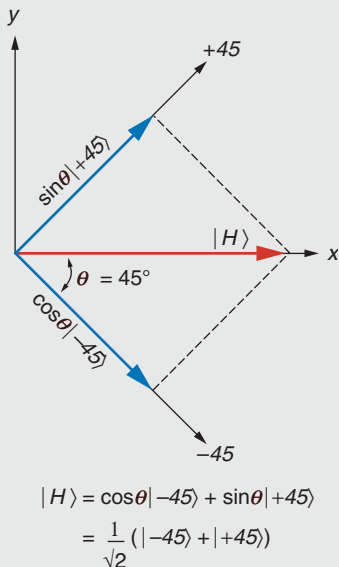


Figure B. Decomposition into Diagonal Basis
A horizontally polarized photon is expressed in terms of the +45/-45 basis.

Photons, Polarizers, and Projection

Our realization of the BB84 protocol uses the polarization state of individual photons to encode bit values. But the key feature that prevents an eavesdropper from detecting the polarizations without being noticed is the use of two nonorthogonal linear polarizations to represent 0 and 1. Rather than preparing a random sequence of horizontally or vertically polarized photons in the quantum states $|H\rangle$ or $|V\rangle$, respectively, Alice (the sender) polarizes photons in the quantum states $|H\rangle$ or $|-45\rangle$ when she wants to send a 0 to Bob (the receiver) and $|V\rangle$ or $|+45\rangle$ when she wants to send a 1.

We can do a simple experiment to demonstrate the quantum mechanical properties of nonorthogonal photons. We need just 3 sheets of linearly polarizing filters, which are readily available from scientific education kits or suppliers. The filter is made from a material that has an intrinsic transmission axis for photons (the polarization axis). As shown in Figure A, if randomly polarized light (for example, sunlight), made up of a large number of photons goes through a linear polarizer with its axis aligned, say, horizontally, the photons that emerge are polarized in the state $|H\rangle$.

We perform our experiment by orienting the first polarizer filter horizontally and holding it up to sunlight. The light intensity decreases by about 50 percent, which indicates that about half the photons get through. We then place a second polarizer behind the first and rotate it until no light passes. At that point, the polarization axes of the two filters are orthogonal to each other, that is, the polarization axis of the second polarizer is in the vertical direction. If we place the third filter between the first two with its polarization axis at -45° to the others, we naively expect no change in the light transmission, but suddenly one eighth of the sunlight gets through the stack, even though the axes of the outer two polarizers are still perpendicular.

These spooky results are a direct consequence of the quantum properties of single photons. A linearly polarized photon is described by a quantum mechanical wave function. Mathematically, it is represented by a “ket” $|\psi\rangle$, which is analogous to an ordinary unit vector in 2-dimensions. Just as a plane vector can be written in terms of two orthogonal plane vectors, we can express $|\psi\rangle$ as a superposition of two orthogonal kets, $|\phi\rangle$ and $|\phi+90\rangle$, in a two-dimensional Hilbert space, with real (as opposed to complex) coefficients. The ket $|\phi\rangle$ represents a photon linearly polarized at the angle ϕ to the horizontal, while $|\phi+90\rangle$ represents a photon polarized at the angle $(\phi+90^\circ)$. The orthogonal kets are a basis for the Hilbert space. We have

$$|\psi\rangle = \cos\theta |\phi\rangle + \sin\theta |\phi+90\rangle, \quad (1)$$

where θ is the angle between $|\phi\rangle$ and $|\psi\rangle$. The coefficients in front of the kets $-\cos\theta$ and $\sin\theta$ —are probability amplitudes. Nature has dictated that the outcome of a measurement of the photon’s polarization state (for example, by transmission through a polarizing filter) is indeterminate—it depends on the basis (the orientation of the polarization axis) used to make the measurement. The probability p that a measurement of $|\psi\rangle$ yields the result $|\phi\rangle$ is given by the expression

$$p = \cos^2\theta, \quad (2)$$

that is, p is the square of the probability amplitude in front of the ket $|\phi\rangle$.

We are now in a position to understand the simple experiment discussed earlier. The polarization axis of the first polarizing filter is set to be horizontal. Equation (1) tells us we can express an incoming photon as a superposition of a ket that is aligned parallel the polarization axis, that is, $|\phi\rangle = |0\rangle \equiv |H\rangle$, and a ket that is orthogonal to the axis, that is, $|\phi+90\rangle = |90\rangle \equiv |V\rangle$. We have

$$|\psi\rangle = \cos\theta |H\rangle + \sin\theta |V\rangle, \quad (3)$$

where the angle θ is now seen to describe the angle between the incoming photon's polarization and the filter's polarization axis. According to Equation (2), the probability that a linearly polarized photon passes through the horizontal polarizer is $p = \cos^2\theta$, that is, the square of the probability amplitude for the state $|H\rangle$. Because photons of all polarizations impinge on the first filter, the amount of light that gets through found by taking the average of p over all angles, that is, $\langle \cos^2\theta \rangle = 1/2$. Half the light makes it through the first filter.

Every photon that makes it through has been projected into the state $|H\rangle$. These photons then interact with the second filter in the stack with polarization axis aligned at $\phi = -45^\circ$. We express the horizontal photon in the diagonal ($-45^\circ/+45^\circ$) basis as (see Figure B):

$$|H\rangle = \cos(45) |-45\rangle + \sin(45) |+45\rangle = 1/\sqrt{2} (|-45\rangle + |+45\rangle). \quad (4)$$

The probability that a photon passes through the second filter is $\cos^2(45) = 1/2$, so $1/4$ of the sunlight makes it through the two filters. The photons that emerge are polarized at -45° . The third filter is aligned vertically ($\phi = 90^\circ$), so we rewrite the ket $|-45\rangle$ in the horizontal/vertical (H/V) basis:

$$|-45\rangle = \cos(-135) |V\rangle + \sin(-135) |-H\rangle = 1/\sqrt{2} (-|V\rangle + |-H\rangle). \quad (5)$$

The probability that a photon passes through the vertical filter is $\cos^2(-135) = 1/2$. Again, half the photons make it through the last filter, so in total one eighth of the sunlight makes it through the stack.

This demonstration of nonorthogonal photon polarizations and polarizers reveals another important property of photons: All information about the initial polarization state is lost as a result of the photon-polarizer interaction. For cryptography, that has an unfortunate implication for someone (Eve) who is trying intercept the encrypted bit stream. Eve can intercept the photons going to Bob, but unless she measures the polarization of those photons in the correct basis, she cannot correlate the results of her measurements with a bit value. With her polarizer set to -45° , she has a probability to detect photons in the state $|-45\rangle$, $|H\rangle$, or $|V\rangle$, corresponding to bit values of 0, 0, and 1. Her measurement does not reveal Alice's bit value, nor does it reveal the original polarization state of the photon. A certain fraction of the photons she sends to Bob (which she must do to cover her tracks) will be in error. Thus, by choosing to send a random sequence of nonorthogonally polarized photons, Alice and Bob assure that Eve cannot attempt to measure the sequence without introducing detectable errors in their QKD protocol.

of those bits. If Eve is to get bit information, she is forced to breach the quantum channel by intercepting the photons and measuring their polarizations. She must then send new photons on to Bob in order to cover her tracks.

But Eve must know the exact state of a photon if she is to send a new one correctly. She cannot, however, make a deterministic measurement of the photon's polarization state because Alice sends photons in two nonorthogonal bases. For example, suppose Eve has a detection apparatus identical to Bob's and she detects a photon in her first detector (bit value of 0) when she measures in the diagonal basis. Did Alice send a photon in the $|H\rangle$, $|V\rangle$, or $|+45\rangle$ state? Eve has no idea because, given her measurement basis, she can detect each of those states. A hit on Eve's detector does not reveal whether Alice sent a 0 or a 1; that information "materializes" only after Alice and Bob compare bases. In fact, Eve can choose any type of detection system or measurement strategy and still be uncertain about the original state of Alice's photon.

One might ask whether Eve can make copies of Alice's photon before making a measurement. Then she could send the original off to Bob, save her string of photons (somehow), and make deterministic polarization measurements after listening to Alice and Bob compare bases. But quantum mechanics prevents Eve from accurately copying an unknown photon. (See the box "The No-Cloning Theorem" on page 79.) She would have to make a deterministic measurement, but that action would inevitably reveal her presence to Alice and Bob.

If she were to guess the polarization state, Eve would have, at best, a 50 percent chance of forwarding the correct one to Bob. But in making her guess, she will necessarily introduce



Figure 3. A Simple Error-Correction Scheme

Error correction removes single-bit errors from the sifted key. A simple scheme involves checking the parity of each byte (8-bit) sequence. The parity of a byte is 0 if the number of 1s in the byte is even or 1 if the number of 1s in the byte is odd. In this case, Alice and Bob start a public conversation to compare the parity of each of their three bytes. Because there is a mismatch, caused by the seventh bit (indicated in red) in the third byte, they try to locate the problem. They must eliminate all errors, or else their keys are unusable. Because the conversation takes place over an open communication line, Eve initially gains information about the parity of the sifted key. That information, however, can be eliminated if Alice and Bob drop some bits from their sequence. Relying on her old information, Eve will not understand anything about the new bit sequence.

errors into Alice and Bob's sifted-key sequence and, hence, increase the BER. When Alice and Bob check their sifted-key sequences for mismatches, they conservatively assume that Eve caused all the errors. They make corrections to those sequences, compute the maximum information Eve could have about the reconciled key, and then use privacy amplification to compress out Eve's possible knowledge about their shared secret strings to substantially

less than one bit. The secret key is truly secret.

Experiments

To date, the three major experiments performed at Los Alamos National Laboratory are free-space, fiber, and entangled-state QKD systems. All of the systems were constructed from readily available pieces of equipment, and we were able to

show that a complete QKD session could be communicated over long distances and still produce a useful secret-bit yield. All three systems use the BB84 protocol.

Here, we describe the free-space and fiber-based experiments. Entangled-state QKD is described on [page 58](#) in the article "Quantum State Entanglement."

Free-Space QKD. In free-space QKD, photons are transmitted through

The No-Cloning Theorem

In 1982, Bill Wootters and Wojciech Zurek applied the linear properties of quantum mechanics to prove that an arbitrary quantum state cannot be cloned. Although their argument is entirely general, we will illustrate the theorem with polarized photons. Suppose we have a perfect cloning device in the initial state $|A_0\rangle$ and an incoming photon in an arbitrary polarization state $|s\rangle$. The device duplicates the photon as follows:

$$|A_0\rangle|s\rangle \rightarrow |A_s\rangle|ss\rangle, \quad (1)$$

where $|A_s\rangle$ is the device final state, which may or may not depend on the polarization of the original photon, and $|ss\rangle$ refers to the state of the electromagnetic field in which there are two photons, each with polarization $|s\rangle$. Suppose that the device can duplicate both the vertical $|V\rangle$ and the horizontal $|H\rangle$ polarization, that is,

$$|A_0\rangle|V\rangle \rightarrow |A_V\rangle|VV\rangle, \text{ and} \quad (2)$$

$$|A_0\rangle|H\rangle \rightarrow |A_H\rangle|HH\rangle. \quad (3)$$

According to quantum mechanics, this transformation should be representable by a linear operator, which means the operator acts independently on each orthogonal state in the Hilbert space. Therefore, if the incoming photon has some arbitrary polarization given by the linear superposition $|s\rangle = \alpha|V\rangle + \beta|H\rangle$, the result of its interaction with the apparatus will be a superposition of Equations (2) and (3):

$$\begin{aligned} |A_0\rangle|s\rangle &= |A_0\rangle (\alpha|V\rangle + \beta|H\rangle) \\ &= \alpha|A_V\rangle|VV\rangle + \beta|A_H\rangle|HH\rangle. \end{aligned} \quad (4)$$

If the apparatus states $|A_V\rangle$ and $|A_H\rangle$ are not identical, the two photons emerging from the apparatus are in a mixed state of polarization; if they are identical, the emerging two photons are in a pure entangled state, $\alpha|VV\rangle + \beta|HH\rangle$. In neither case does the apparatus produce a final state $|ss\rangle$ consisting of two completely independent photons, each in the polarization state $\alpha|V\rangle + \beta|H\rangle$:

$$\begin{aligned} |ss\rangle &= (\alpha|V\rangle + \beta|H\rangle) (\alpha|V\rangle + \beta|H\rangle) \\ &= \alpha^2|VV\rangle + \alpha\beta|VH\rangle + \beta\alpha|HV\rangle + \beta^2|HH\rangle. \end{aligned} \quad (5)$$

Linearity, therefore, rules out the existence of a device that could faithfully clone a photon in an arbitrary polarization state.

open air. The protocol uses polarization states, as previously described, because the atmosphere preserves polarization over a wide range of photon wavelengths (including the full range of visible and infrared light). The major difficulty is detecting the

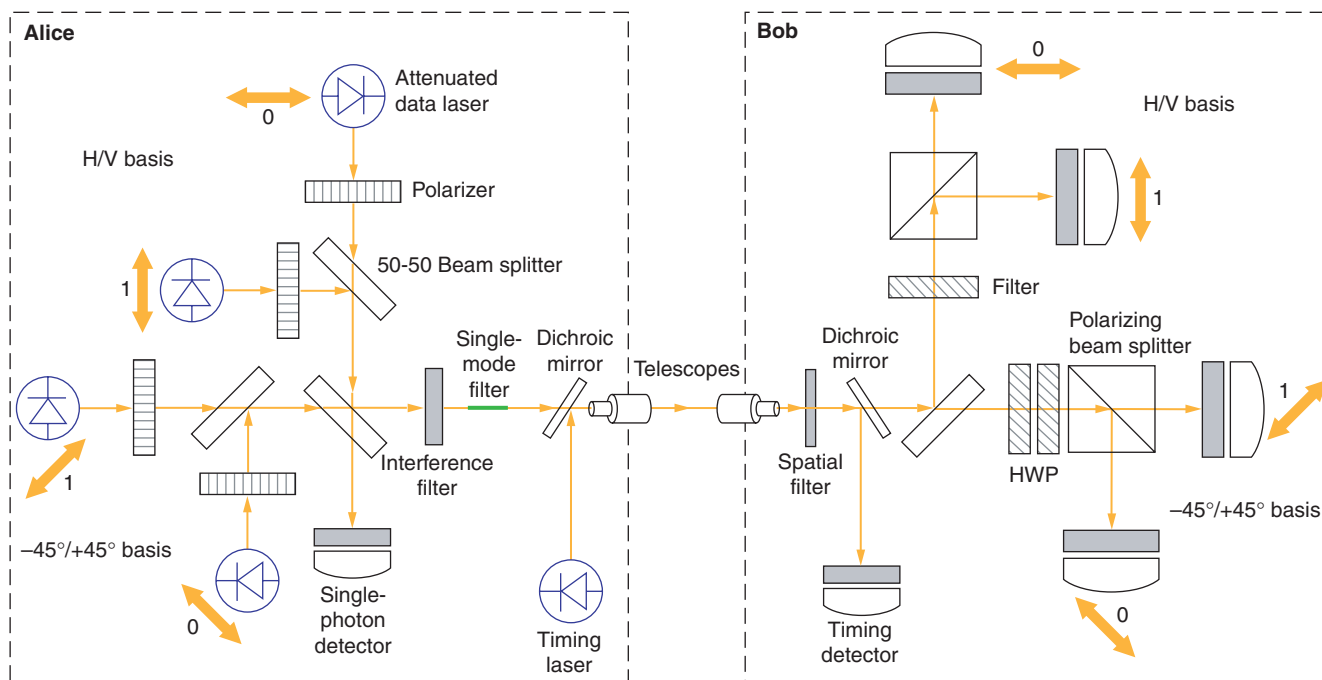
single QKD photons from within the enormous background of daytime photons, namely, $\geq 10^{10}$ background photons per centimeter squared, per second, per angstrom, per steradian ($\gamma/\text{cm}^2/\text{s}/\text{\AA}/\text{sr}$). This problem exists even at night because the background

from, say, moonlight or the light of urban areas is still much larger than the QKD signal. A second difficulty is dealing with losses due to atmospheric distortions. We are able to overcome both of these problems and can distinguish the QKD photons from background photons by using interference filters that transmit only photons of a specific wavelength, by carefully limiting the field of view, and by using a clever trick. The free-space QKD system is shown in Figure 4.

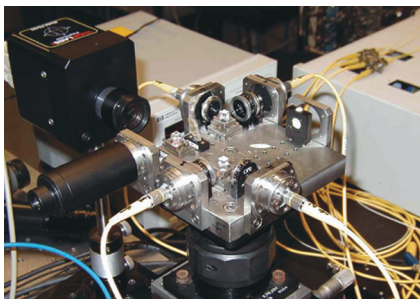
Alice and Bob have identical copies of the interference filters, which allow Alice to send photons at a selected wavelength and Bob to receive photons only at that wavelength. The preferred wavelength is about 772 nanometers, which is in the infrared and just outside the normal range of vision. The atmosphere is highly transmitting for light of this color, and single-photon detectors with good quantum efficiency at this wavelength are readily available. Furthermore, polarization selection and control components and diode lasers that produce the desired wavelength are all easily obtained.

A receiver telescope with a narrow field of view helps limit unwanted photons. Behind the telescope is a spatial filter that passes photons coming from a precise location (Alice's) while excluding all the others. The telescope must be employed with care, however. As anyone who has ever looked at the twinkling stars knows, the atmosphere can make a source of light appear to move. The magnitude of the movement varies considerably with the time of day, the weather, and the local terrain. If not accounted for, the atmosphere could cause Alice to shift rapidly in and out of Bob's field of view. Over short distances, these atmospheric distortions are not a serious problem. Over long distances, Alice

(a) Conceptual Diagram



(b) Alice's Optics Table



(c) Alice's Electronics



(d) Bob's System

**Figure 4. Free-Space QKD**

(a) In the BB84 protocol, Alice (the sender) encodes bits in the polarization states of single photons either as $0 = |H\rangle$ and $1 = |V\rangle$ or as $0 = |-45^\circ\rangle$ and $1 = |+45^\circ\rangle$. The data stream begins with a bright output pulse from the timing laser, which sets the timing of the pulse. A few nanoseconds later, one of the four data lasers ($\lambda = 772 \text{ nm}$) fires. Each data laser has its own attenuator, focusing optics, and polarizer. Each laser outputs a uniform pulse of the desired brightness in one of the four polarization states. The output of all four data lasers is combined by a series of beam splitters, which have been carefully arranged so that the distances between the lasers and output optics are the same (therefore eliminating any timing differences between the pulses). The final beam

splitter either directs the photons to a detector that monitors the average number of photons per laser pulse or sends the polarized photons through a narrow-pass interference filter (to remove any frequency differences) and a single mode fiber (to eliminate any spatial mode differences). The photons that pass through Alice's telescope are identical in every respect except for polarization. Bob (the receiver) uses spatial filtering, time-domain filtering, and wavelength selection to pick out Alice's photons from background. His telescope, with a field of view that is nominally 45 arc seconds (or 220 microradians), acts as a spatial filter that allows only photons from Alice's location to pass. The photons then pass through an interference filter (wavelength selection)

that is matched to the one in Alice's transmitter. Photons are sent to a 50-50 beam splitter, which acts as a basis selector by randomly directing a photon to one of the two measurement stations. Each station consists of a polarizing beam splitter and two single-photon detectors. A half-wave plate (HWP) rotates the photon's polarization before the $-45^\circ/+45^\circ$ station. A detector must fire within a set period following detection of the bright timing pulse (time-domain filtering). (b) Alice's compact optics table and (c) electronics are shown here. (d) Bob's telescope peers out from the door of the mobile trailer containing all his electronics and optical systems. Bob (and Alice) can be easily transported to different sites. Moreover, one person can operate the system.

corrects for atmospheric variations by observing Bob's beacon laser and is thus able to rapidly vary the point to which she sends the photons.

Finally, the clever trick is to send a bright laser pulse from Alice to Bob just before a single photon is sent so there is a known delay between the photon and the bright pulse. Bob accepts only photons that enter the system approximately 1 nanosecond after the bright pulse. This time-domain filtering greatly limits the possibility of a back-ground photon being detected instead of a QKD photon. This system of multiple filtering techniques works so well that single QKD photons can be distinguished from back-ground even in daylight.

One issue complicating the free-space system (as well as the other systems described below) is that the photon sources are actually attenuated laser diodes that produce weak laser pulses instead of true single photons. (Single-photon sources are currently too large and exotic for systems intended for use in the field.) The number of photons in a weak laser pulse is governed by Poisson statistics, and the number of photons in each pulse varies. The probability $P(n)$ that a pulse will contain n photons is,

$$P(n) = \frac{e^{-\mu} \mu^n}{n!}, \quad (4)$$

where μ is the average number of photons per pulse. If $\mu = 1$, there is roughly a 37 percent chance that a pulse will contain no photons, 37 percent that it will contain one photon, and 26 percent that the pulse will contain more than one photon.

By adjusting the attenuation, Alice can choose a specific value of μ . If she chooses a relatively high μ , say, above 1 photon per pulse, each time more than one photon is sent, it must be assumed that a clever eaves-

dropper would be able to detect and measure the extra photons. A great deal of privacy amplification—concomitant with a large consumption of reconciled bits—is needed to keep the system secure, so overall, the secret bit yield decreases. If μ is too small, say, 0.05, then most of the time Alice is sending nothing over the quantum channel and experimental errors (such as back-ground light getting into the receiver, dark counts in detectors, or even the actions of an eavesdropper) may dominate. Again, the secret-bit yield decreases. The choice of μ is therefore an important free parameter at Alice's disposal.

Our experiments have shown that the secret-bit yield depends strongly on atmospheric conditions. Turbulence along the optical path between Alice and Bob, for example, affects the transmission efficiency. To help show trends in the data, we construct a pseudo signal-to-noise ratio, η/C , where η is the transmission efficiency (obtained by dividing the number of sifted bits by μ) and C is the number of back-ground photons detected by Bob.

Figure 5 shows data from a free-space QKD experiment that ran successfully at a 10-kilometer separation in daylight. The open communication channel was a wireless Ethernet. During the numerous experimental runs, Alice would send 10^6 laser pulses over a 1-second period. The value of μ was typically set between 0.1 and 0.8.

The experimental run labeled "Sample" in Figure 5 is a typical example. Approximately 22 percent of the pulses had a single photon ($\mu = 0.29$). After comparing Alice and Bob's bases, we constructed a sifted key of 651 bits. Following error correction, calculation of the BER, and privacy amplification, we obtained a secret key consisting of 264 bits, which is sufficient for the

new AES. Note that the secret-bit yield can be substantially higher at night (high η/C), because the back-ground is reduced.

Our free-space system is a preliminary prototype for a system that could be flown on a spacecraft. Because the atmosphere has an effective thickness of only a few kilometers if one were to look straight up, our results are a good indicator of the feasibility of ground-to-satellite free-space QKD.

Fiber-Based QKD. The polarization state of a photon is not preserved in conventional optical fibers. That is why another physical property that could express the desired quantum mechanical properties for QKD had to be found in order to implement a fiber-based system. The solution was to have a photon interfere with itself after it travels down two paths of a twin Mach-Zehnder interferometer setup.

The concepts underlying the fiber-based QKD scheme are illustrated in Figure 6. Briefly, quantum mechanics tells us that a single photon entering a Mach-Zehnder interferometer behaves as if it has taken both paths through the instrument. The entrance beam splitter places the photon in a quantum mechanical superposition, with a component that describes a photon traversing the upper path and a component that describes the photon traversing the lower path. The two components have a definite phase relationship and can interfere with each other.

As seen in the figure, Alice can introduce a phase shift ϕ_A to the photon on one arm of the interferometer, while Bob can introduce a phase shift ϕ_B on the other. Depending on the phases set by both Alice and Bob, the interference at the exit beam splitter is such that the photon has a definite probability

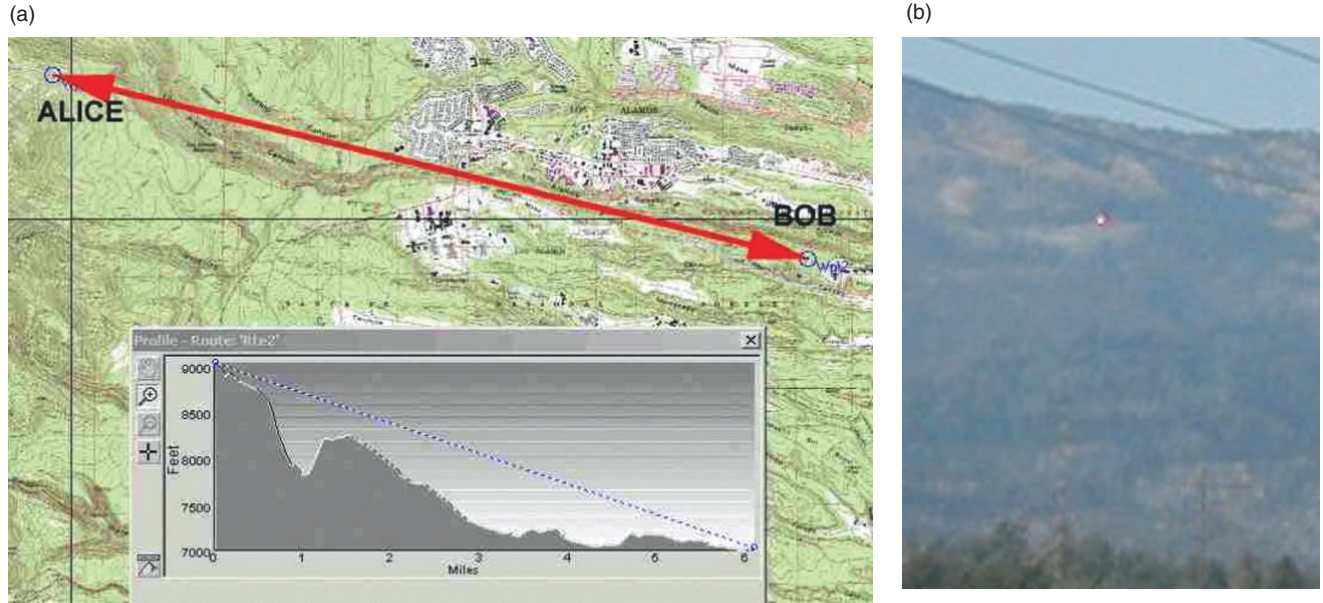


Figure 5. Data from a 10-km Free-Space QKD Experiment

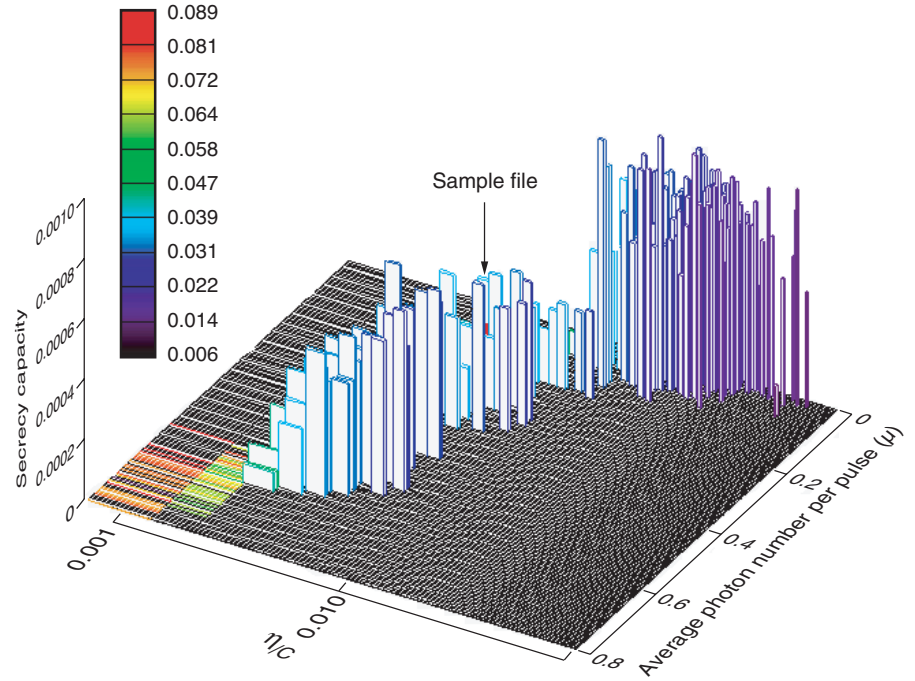
(a) Alice was located halfway up Pajarito Mountain, in New Mexico, while Bob was 10 km away, at a Los Alamos lab site. (b) The bright red dot near the center of the picture is a spotting laser sent through Alice's telescope. It was used to optically align the transmitter and receiver for the quantum channel. (c) Data from the experiment show the dependence of the secret-bit yield (normalized to the number of sifted bits) on the average number of photons per pulse μ and on the pseudo signal-to-noise ratio η/C (discussed in the text). Each vertical column corresponds to an experimental run in which Alice sent 10^6 polarized photons in 1 s. The flat, black regions of the graph are areas for which no data are available. With favorable atmospheric conditions or low background (high η/C), we can run at lower μ values and still obtain a high bit yield. Poorer conditions (low η/C) require higher μ values and result in a lesser yield.

to hit either of two detectors. The probability P_U that the photon hits the upper detector is given by

$$P_U = \sin^2\left(\frac{\phi_A - \phi_B}{2}\right), \quad (5)$$

whereas the probability P_L that the photon hits the lower detector

(c) Sifted-Bit Error Rate (r)



is given by

$$P_L = \cos^2\left(\frac{\phi_A - \phi_B}{2}\right) \quad (6)$$

We make use of these relations to implement the BB84 protocol. Alice chooses at random between two bases, X and Y. If she chooses the X-basis, then

for a bit value of 0 or 1, she sets $\phi_A = 0^\circ$ or 180° , respectively. If Alice chooses the Y-basis, then she chooses $\phi_A = 90^\circ$ or 270° for bit values of 0 or 1, respectively. At his end, Bob sets his phase angle ϕ_B to 0° if he is in the X-basis and to 90° if he is in the Y-basis.

Table II summarizes Alice and Bob's choices and shows the value of the probabilities P_U and P_L , given the

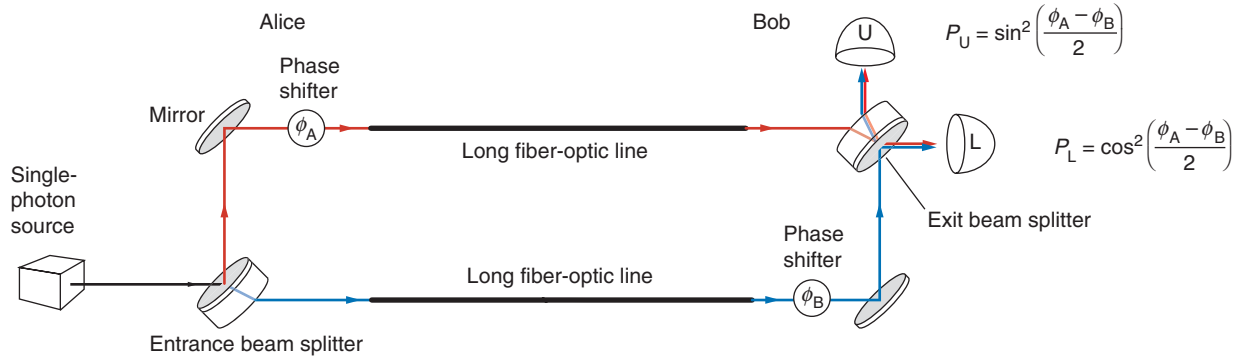


Figure 6. Mach-Zehnder Interferometer and Fiber-Based QKD Concept

In a Mach-Zehnder interferometer, a photon is placed in a superposition of two states by the entrance beam splitter. It travels down both arms simultaneously, and interferes with itself at the exit beam splitter. In the conceptual fiber-based QKD system illustrated here, a phase shifter is placed in each arm of the interferometer. Alice randomly chooses a bit value and a basis and sets the angle of her phase shifter according to her choices (see Table II below). Bob sets the

angle of his phase shifter according to his basis choice. The table shows the probability that Bob detects a photon in a given detector. When Alice and Bob use the same basis for sending and measuring, a hit in Bob's lower detector means that Alice sent a bit value of 0, whereas a hit on the upper detector means she sent a 1. Because there is no such correlation when Alice and Bob use different bases, those bit values are discarded.

Table II. Fiber-Based QKD

Sender (Alice)			Receiver (Bob)			Action		
Basis	Bit	Phase ϕ_A (°)	Basis	Phase ϕ_B (°)	Probability (%)		Bit	
					P_L	P_U		
X	0	0	X	0	100	0	0	Keep bit
X	1	180	X	0	0	100	1	Keep bit
X	0	0	Y	90	50	50	0 or 1	Discard bit
X	1	180	Y	90	50	50	0 or 1	Discard bit
Y	0	90	X	0	50	50	0 or 1	Discard bit
Y	1	270	X	0	50	50	0 or 1	Discard bit
Y	0	90	Y	90	100	0	0	Keep bit
Y	1	270	Y	90	0	100	1	Keep bit

various combinations of ϕ_A and ϕ_B . Because we are implementing BB84, Table II is essentially the same as Table I. When Alice and Bob choose the same basis, a photon representing Alice's 1 always goes to the upper detector, and a photon representing her 0 always goes to the lower. If Alice and Bob use different bases, the photon has equal probability to emerge from either port, and Bob has no information about what bit value Alice has sent. At the end of the session, Bob calls Alice on the open

communications line, and the two compare which bases they used for each photon. They keep the bit values when the bases agree and discard the other bits.

In the scheme discussed above, a single Mach-Zehnder interferometer stretches between Alice and Bob. In practice, that is a bad idea. The photon needs to maintain phase coherence as it propagates down the two optical fibers that make up the long arms of the interferometer. Photons often experience random phase shifts as

they go through long fiber-optic cables, and because the shifts in one arm are independent of those in the other, the interference condition at the exit beam splitter changes in a random fashion. Furthermore, having two dedicated fibers would be expensive to operate in the real world.

A better idea is for Alice and Bob each to have a Mach-Zehnder interferometer, with the two connected by a single long fiber—see Figure 7.

Each interferometer is modified to have a long arm and a short arm, and

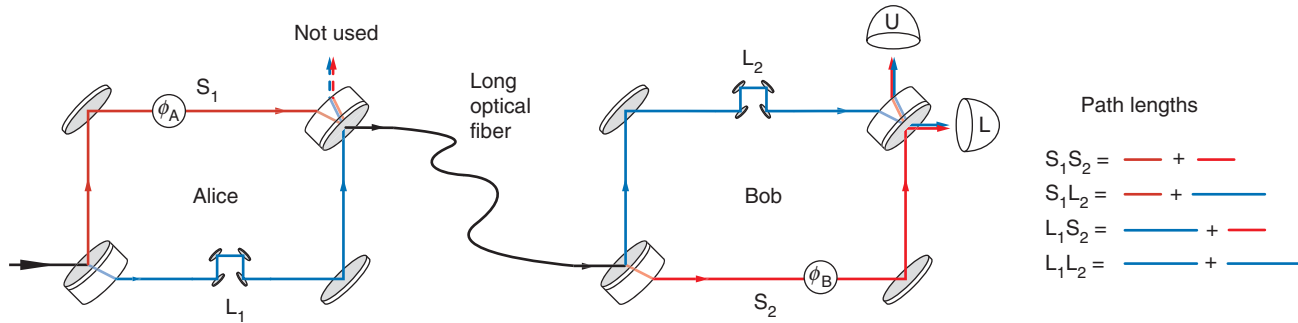


Figure 7. Implementation of Fiber-Based QKD

Our fiber-based QKD system uses two modified interferometers connected by a single, long optical fiber. Each interferometer has a long (L) arm and a short (S) arm. In going from Alice's entrance beam splitter to Bob's exit beam splitter, the photon can take paths S_1S_2 , L_1L_2 , S_1L_2 , and L_1S_2 . The latter two paths have the same length, and the photon traveling them can maintain phase coherence and interfere with itself. The protocol then works as described in Figure 6.

the path length differences between the two arms are greater than the coherence length of the photon. There is no interference as the photon leaves Alice's instrument. But of the four possible paths through the entire system (refer to the figure), the two designated as S_1L_2 and L_1S_2 are of equal length (to within the phase coherence length of the photon). A photon that travels down those two paths interferes with itself at Bob's exit beam splitter. The system therefore behaves as if it were a single instrument. Alice and Bob are still free to vary the phase on one arm of their interferometers, as needed, to carry out the protocol.

Our system transmits bits through 48 kilometers of fiber. As in the free-space experiments, Alice first sends a bright pulse to trigger the detectors and to limit background interference. Single photons are sent at 1310 nanometers, and the bright timing pulse is at 1550 nanometers. The secret-bit yield is lower than that obtained in the free-space experiment.

Summary

Quantum cryptography can enable secure transmission of sensitive, proprietary, or national security information across a metropolitan area or

corporate campus and provide the long-term security guarantees such data require. It is the only technology that will be secure no matter what technology an adversary develops in the future. Furthermore, it raises the stakes for eavesdroppers because they must perform risky, active attacks against a system. Currently, a public-key encrypted system can be attacked through passive, standoff monitoring.

Because of the inherent advantages of quantum cryptography, we can envision a future in which a QKD system provides secure communications in metropolitan areas between banks, between off-site stock-trading centers and central stock exchanges, between corporate offices, and between offices and broadband data networks. Money transfers between banks now amount to over \$2 trillion per day worldwide and well justify the expense of implementing QKD systems. Optical wireless "last-mile" communications systems could even provide broadband access to most homes.

By combining theoretical analyses with innovative experimental advances, the Los Alamos quantum cryptography team has already demonstrated the practicality of free-space quantum cryptography in a series of record-setting experiments. In 1996, the team demonstrated

atmospheric quantum-key transmission at night, quickly followed by a record-setting 0.5-kilometer point-to-point transmission in full daylight, then a 1.6-, and finally a 10-kilometer transmission. The world record for the longest QKD distribution in fiber—48 kilometers—was also held by the Los Alamos team for many years. Several of the first demonstrations of entanglement-based QKD have also been performed at the Laboratory.

In the near future, the free-space quantum cryptography system could provide secure satellite communications—using a low-orbit satellite—between cities anywhere in the world. When deployed on a spacecraft, our system can be used to generate cryptographic keys between any two users who are anywhere on the planet and can view that spacecraft. Each user would individually generate a key with the spacecraft. The second user would then be instructed to change specific bits so that the two users' keys would match. Because the spacecraft only needs to instruct the user which bits to change, and can do so without revealing any bit values, this is a secure key-generation methodology.

On a more philosophical note, the challenging demands of cryptography have already produced a huge growth in research into the foundations of

quantum mechanics. Fundamental concepts that were previously thought to be testable only in thought experiments have been subjected to experimental verification. Many concepts, such as entanglement, that have been almost completely neglected since the early days of quantum physics have been explored and realized. This trend will continue, and we will find out to what extent the creation and control of “mesoscopic” quantum systems, that is, the netherworld between single-particle behavior and collective-particle behavior, can be performed. This research may help elucidate the puzzling transition between the quantum and classical regime. The development of quantum technology will open up other applications of quantum physics, such as quantum-enhanced sensors and improvements to atomic clocks and satellite navigation systems. Whether or not quantum cryptography becomes a widely adopted technology, we are in for an interesting next decade. ■

Acknowledgments

The Quantum Cryptography team combines the talents of numerous scientists and engineers, including those of Kevin P. McCabe, George L. Morgan, Michael J. Pigue, Steven A. Storms, Paul A. Montano, James T. Thrasher, and especially Charles G. Peterson. The authors wish to thank Derek Derkacs for technical support. We gratefully acknowledge support for the 10-kilometer free-space experiment from the National Reconnaissance Office Director's Innovation Initiative program, administered by Col. John Comtois and Peter Hendrickson.

Further Reading

- Bennett, C. H. 1992. Quantum Cryptography: Uncertainty in the Service of Privacy. *Science* **257** (5071): 752.
- Bennett, C. H., and G. Brassard. 1984. Quantum Cryptography: Public-Key Distribution and Coin Tossing. In *Proceedings of the IEEE International Conference on Computers, Systems, and Signal Processing, Bangalore, India, 1984*. p. 175. New York: IEEE.
- Bennett, C. H., G. Brassard, C. Crépeau, and U. M. Maurer. 1995. Generalized Privacy Amplification. *IEEE Trans. Inf. Theory* **41** (6): 1915.
- Bennett, C. H., G. Brassard, and A. K. Ekert. 1992. Quantum Cryptography. *Sci. Am.* **267** (4): 50.
- Hughes, R. J., D. G. L. Morgan, and C. G. Peterson. 2000. Quantum Key Distribution over a 48-km Optical Fiber Network. *J. Mod. Opt.* **47**: 533.
- Hughes, R. J., J. E. Nordholt, D. Derkacs, and C. G. Peterson. 2002. Practical Free-Space Quantum Key distribution over 10 km in Daylight and at Night. *New J. Phys.* **4**: 43. [Online]: <http://www.njp.org>
- Hughes, R. J., W. T. Buttler, P. G. Kwiat, S. K. Lamoreaux, G. L. Morgan, J. E. Nordholt, and C. G. Peterson. 2000. Free-Space Quantum Key Distribution in Daylight. *J. Mod. Opt.* **47**: 549.
- Hughes, R. J., W. T. Buttler, P. G. Kwiat, S. K. Lamoreaux, G. L. Morgan, J. E. Nordholt, and C. G. Peterson. 2000. Quantum Cryptography for Secure Satellite Communications. In *2000 IEEE Aerospace Conference Proceedings*. p. 191. New York: IEEE.
- Hughes, R., and J. Nordholt. 1999. Quantum Cryptography Takes to the Air. *Phys. World* **12** (5): 31.
- Hughes, R. J., W. T. Buttler, P. G. Kwiat, S. K. Lamoreaux, G. G. Luther, G. L. Morgan, J. E. Nordholt, and C. G. Peterson. 1999. Quantum Cryptography for Secure Free-Space Communications. *Proc. SPIE-Int. Soc. Opt. Eng.* **3615**: 98.
- Nordholt, J. E., R. J. Hughes, G. L. Morgan, C. G. Peterson, and C. C. Wipf. 2002. Present and Future Free-Space Quantum Key Distribution. *Proc. SPIE-Int. Soc. Opt. Eng.* **4635**: 116.
- Schneier, B. 1995. *Applied Cryptography: Protocols, Algorithms Source Code* in C. New York: John Wiley & Sons.
- Singh, S. 1999. *The Code Book: The Evolution of Secrecy from Mary, Queen of Scots to Quantum Cryptography*. New York: Doubleday.
- Wootters, W. K., and W. H. Zurek. 1982. A Single Quantum Cannot be Cloned. *Nature* **299**: 802.

Jane E. (Beth) Nordholt has broad experience in quantum-key distribution, experimental astrophysics, high-energy physics, computing, and space plasma physics. Currently a technical



project leader at Los Alamos National Laboratory, she is the coinventor for the free-space quantum key distribution project and holds several patents on quantum-key distribution and spacecraft instrumentation. Beth received

four NASA group achievement awards and two Los Alamos Distinguished Performance Awards. In 2001, she received an R&D 100 Award for her work on free-space quantum cryptography from the *Research and Development* magazine. Her interests include quantum cryptography, quantum communications, quantum metrology, the composition of planetary magnetospheres, planetary science, and advanced instrumentation.

Richard J. Hughes is a Laboratory Fellow and Quantum Information Science team leader in the Neutron Science and Technology Group of the Physics Division at Los Alamos National Laboratory. He is the principal investigator for several projects in quantum computation and quantum cryptography. Richard obtained his Ph.D. in theoretical elementary particle physics from the University of Liverpool and held research positions at Oxford University and The Queen's College, Oxford; California Institute of Technology; and CERN, the European Center for Nuclear Research. He was a distinguished visiting scientist at Oxford University and the University of Oslo. Richard was awarded the Los Alamos Fellows Prize for his work on quantum information science; he was twice awarded Los Alamos Distinguished Performance Awards for his quantum cryptography research; and he was cowinner of an R&D 100 Award for the entry “Free-Space Quantum Cryptography.” He became a Fellow of the American Physical Society in 1999. He has authored over 100 scientific papers on quantum field theory, the foundations of quantum mechanics, quantum cryptography, and quantum computation. In his spare time, Richard enjoys ultramarathon trail running over distances of up to 100 miles.

